



DOI: 10.22363/2313-2299-2025-16-4-1272-1287
EDN: NRNDSC

Research Article / Научная статья

Features in Lexicography for Arabic Dialects: Multimedia Dictionary of the Syrian Dialect

Natalia V. Novospasskaya , Aida D. Haddad  

RUDN University, Moscow, Russian Federation
 khaddad_a@pfur.ru

Abstract. The linguistic situation in Arab countries is characterized by the presence of several variants of the Arabic language, which interact constantly, complicating the establishment of a lexical corpus for each dialect. This paper analyzes existing approaches to the study of Arabic dialects, identifies shortcomings in the methodology of dialect analysis, and discusses directions for comparative research. It establishes the growing research interest in dialects, highlights the lack of systematic analysis of dialectal vocabulary in Arabic, and notes the interest in the phonetic and grammatical levels of analysis of specific dialects, as well as the lexical level in comparative studies. The authors propose an overview of existing digital solutions to describe and analyze Arabic dialects, highlighting the need to find new lexicographic formats and sources of dialectal material that can help capture the lexical composition and phonetic and grammatical characteristics of dialectal units. They also propose a new approach to organizing a dialectal dictionary and present a project for a multimedia dictionary of the Syrian dialect of Arabic.

Keywords: dialectology, Arabic language, dictionary

Authors' contribution: the authors contributed equally to this research and preparation of the article text.

Conflicts of interest: the authors declare no conflicts of interest.

Financing: the publication has been supported by the RUDN University Scientific Projects Grant System, project No 050225-0-000 “Multimedia dictionary of specialized vocabulary”.

Article history: received: 23.07.2025; accepted: 15.08.2025.

For citation: Novospasskaya, N.V., & Haddad, A.D. (2025). Features in Lexicography for Arabic Dialects: Multimedia Dictionary of the Syrian Dialect. *RUDN Journal of Language Studies, Semiotics and Semantics*, 16(4), 1272–1287. <https://doi.org/10.22363/2313-2299-2025-16-4-1272-1287>

© Novospasskaya N.V., Haddad A.D., 2025



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Особенности лексикографирования арабских диалектов: мультимедийный словарь сирийского диалекта

Н.В. Новоспасская , А.Д. Хаддад  

Российский университет дружбы народов, Москва, Российская Федерация
 khaddad_a@pfur.ru

Аннотация. Языковая ситуация в арабских странах характеризуется наличием нескольких вариантов арабского языка, находящихся в постоянном взаимодействии, что затрудняет установление лексического корпуса каждого диалекта. В работе анализируются существующие подходы к изучению арабских диалектов, выявляются недостатки в методологии анализа диалектов арабского языка, рассматриваются направления сопоставительных исследований. Установлен рост исследовательского интереса к диалектам, зафиксирована проблема отсутствия системности анализа диалектной лексики арабского языка, отмечен интерес к фонетическому и грамматическому уровням анализа конкретных диалектов и к лексическому уровню в сопоставительном изучении. Авторами предложен обзор существующих цифровых решений для описания и анализа диалектов арабского языка, отмечена необходимость в нахождении новых лексикографических форматов и источников диалектного материала, способствующих фиксации лексического состава и фонетических, грамматических характеристик диалектных единиц. Авторы предлагают новый подход к организации диалектного словаря и представляют проект мультимедийного словаря сирийского диалекта арабского языка.

Ключевые слова: диалектология, арабский язык, словарь

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

Финансирование: публикация подготовлена в рамках грантовой поддержки научных проектов РУДН 050225-0-000 «Мультимодальный словарь специальной лексики».

История статьи: дата поступления: 23.07.2025; дата приема в печать: 15.08.2025.

Для цитирования: *Novospasskaya N.V., Haddad A.D. (2025). Features in Lexicography for Arabic Dialects: Multimedia Dictionary of the Syrian Dialect // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 4. С. 1272–1287. <https://doi.org/10.22363/2313-2299-2025-16-4-1272-1287>*

Introduction

The linguistic situation in the Arab countries is characterized by linguists as the presence of “sociolinguistic phenomena such as bilingualism, multilingualism, and diglossia.” Researchers note that multilingualism (or polylingualism) is a striking sociolinguistic feature in the Arab world, which presents an extensive and interesting field of study for linguists, sociologists, and cultural scholars [1. P. 52]. This multilingual context complicates the study of Arabic dialects in several ways:

1. Determining the lexical corpus of a dialect is difficult, as Arabic dialects closely interact with the codified literary Arabic language, leading to partial overlap in the lexical systems and making it challenging to distinguish dialectal from common colloquial vocabulary.

2. Sociolinguistic factors must be considered, as the prestigious status of the literary language and the lower social status of dialects can result in code-switching by native Arabic speakers, who may adapt their speech to the communication situation by switching between dialect and literary language. Reem Bassiouney argues that Arabic speakers perceive all language varieties as “Arabic,” as they originate from regions where Modern Standard Arabic is the official language, and they do not distinguish between classical Arabic (CA) and modern standard Arabic (MSA) [2].

3. There is no consensus on the forms of the Arabic language, with some identifying three forms (the language of the Quran, literary Arabic, and dialectal forms), while others distinguish between literary Arabic (including both classical and modern standard forms) and colloquial/dialectal varieties [3–5].

4. There is significant regional variability in Arabic dialects, forming a linguistic continuum, making it difficult to unambiguously classify a linguistic form as either literary Arabic or a specific dialect.

5. In this context, we can distinguish three main varieties of the Arabic language: literary Arabic (the language of the Quran and modern literary Arabic), colloquial Arabic, and regional dialects. The differentiation between the literary and dialectal forms is primarily based on their domains of use, with literary Arabic being used in official, religious, and written spheres, while dialects are used in everyday oral communication. However, dialects can also be used in some formal situations and literary works, blurring the line between “high” literary and “low” dialectal language, as seen in the emergence of the poetic trend شعر بالمحكية [shi’ir bel’miḥq’ie] (poems in colloquial language).

Literary Arabic, in its classical and modern standard forms, is the codified and official language of all Arab countries, functioning as the language of writing, education, media, religion, and legal documents. Colloquial Arabic, on the other hand, is an unstandardized version that arises from the use of literary lexical units and constructions with minimal phonetic or grammatical changes. This colloquial variety is understood by all Arabic speakers and serves a primarily communicative function, allowing adaptation of speech for interaction between speakers of different dialects.

These three varieties of Arabic are in complex interaction, with literary Arabic influencing the colloquial and dialectal forms, while the dialects remain the most changeable variants, as they do not strive to comply with literary norms like the colloquial variety.

Materials and methods of the research

The research comprised several stages, one of which involved analyzing existing trends in the study of Arabic dialects and reviewing contemporary scholarly works. The materials for this research included scientific articles and monographs focused on the Syrian dialect of Arabic, as well as comparative studies of various Arabic dialects. The final stage entailed the development of a multimedia

dialect dictionary for the Syrian dialect of Arabic, incorporating elements of the lexical network of the Arabic language¹.

Research results

The study showed that

1) at present, there is an increase in research interest in the dialects of the Arabic language and only conditionally can we talk about the presence of a developed methodology for analyzing dialectal linguistic units of all levels, since existing works demonstrate the use of various approaches to studying certain phenomena / lexemes, and do not offer a systemic analysis of the structure of the dialect in question (or one of its levels) as a whole;

2) there is a significantly large layer of research devoted to recording and identifying patterns in dialects and differences from the literary language at the phonetic level; at the grammatical level, research is focused primarily on the word-formation development of dialect forms and the analysis of the development of the verb system of dialects; principles for describing the syntactic level of Arabic dialects are developed; primary attention is paid to individual aspects of the lexical level of Arabic dialects;

3) comparative works largely reflect the existing methodological problems of dialect studies, allow us to see the heterogeneity of approaches to the description of dialect systems, the lack of generally accepted criteria for choosing dialects for comparison and establishing genealogical links, and also show the inadequacy of dialect research, which is expressed in the lack of the ability to conduct a full-fledged comparative historical analysis and identify general trends and patterns in the development of Arabic dialect speech;

4) the introduction of new tools in work with dialects allows us to expand the range of research tasks, in particular, the development of electronic resources for the “normalization” of dialects by translating them into literary form;

5) in order to continue the study and preservation of the linguistic material of Arabic dialects and the culture of their speakers, it is necessary to create electronic, including multimedia, dictionaries.

Discussion

Level studies of the Syrian dialect of Arabic

Since ancient times, the location of Syria and Greater Syria (Bilad al-Sham — the union of Syria, Lebanon, Palestine and Jordan with the capital in Damascus) was strategically important. This region, which includes Syria, Lebanon, Palestine, and Jordan with its capital in Damascus, served as a link between three continents,

¹ Lexical Network — a graph-based model of a language’s vocabulary, within which different types of relations between words (equivalence, antonymy, motivating-motivated) are shown using graphically distinct lines (Authors. — *A.H., N.N.*)

Aleppo, Damascus, Latakia, Palmyra, Dayr ez-Zor were cities of the Great Silk Road. “What other country, except Syria, can claim to have seen with its own eyes such a brilliant galaxy of warriors and conquerors of the world, starting with Thutmose, Nebuchadnezzar, Alexander the Great and Julius Caesar, continuing with Khalid ibn al-Walid, Saladin and Baybars and ending with Napoleon?”².

The Arabic language arrived in the region with the advent of Islam in the 7th century in two forms: literary Arabic and colloquial Arabic. Before that, it was primarily used orally by some tribes. LA was almost not subject to change, but the colloquial version was assimilated with the Aramaic and Syriac languages that were in the territory at that time. Literature in the Syriac language actively developed in Syria even in the first centuries of Arab domination. The significant influence of the Syriac language on Arabic can be observed in the study of Arabic etymology. Shapes of Arabic graphemes, the order of letters and their names also indicate a close connection between Arabic and Syriac languages. Another indicator of this influence is that both languages share the same graphemes that are not connected to the following letters when written: ر [r], د [d]. In many Arab countries, the names of settlements and regions of Syriac origin have been preserved: كفر [kafr], دير [daj'r], معرة [maara] (Table 1).

Table 1. Toponyms of Syriac Origin

معناها بالسرياني	الكلمة المتداولة	معناها بالسرياني	الكلمة المتداولة
مغارة الشحم	معارات الأثارب	المغائر	معرثا
مغارة خان طومان	مغارة خان طومان	مغارة المسلمية	مغارة المسلمية
مغارة الشيخ	مغارة الشلف	مغارة الماء	معرمايا
مغارة الدبس	مغارة دبسا	المغارة الشمالية	مغارة الشمالية
مغارة المتانيين	مغارة مصرين	المغارة العلوية	مغارة عليا
مغارة الفود أو المحراك	مغارة متائر	المغارة المحرمة	مغارة حرمة
مغارة اشقي السهام	مغارة شمارين	مغارة الزيت	مغارة زيتا
مغارة الشمس	مغارة شمشا	المغارة ذات الرائحة الكريهة	مغارة شمسين
مغارة صغيرة	مغارة معرونة	مغارة الاسوار	مغارة شورين
مغارة سوداء	مغارة مشحور	مدخل	مغارة معلولة

Source: compiled by Aida D. Haddad & Natalia V. Novospasskaya.

One of the indicators of the great influence of the One indicator of the significant influence of the Syriac language on Arabic in Syria is its presence in three localities: Jabdin and Bahaa, where the inhabitants are Muslims, and Maaloula

² Hitti, P. (2022). History of Syria: History of Syria: Including Lebanon and Palestine; translated from English by T.M. Shulikova. Moscow: Tsentrpoligraf. P. 11.

(Damascus Province), where both Muslims and Christians reside. Currently, the inhabitants of Maaloula speak Aramaic, Christian services are conducted in Aramaic, Greek and Arabic. The population of the city is about 3,000 people. The vocabulary of Syrian dialects reflects the period of the country's entry into the Ottoman Empire, which began in 1517 when Syria was conquered by Sultan Selim I. The borrowing of vocabulary from Turkish peaked during this period. At the beginning of the 19th century, Turkish national parties implemented a policy of Turkification, aiming for the linguistic assimilation of non-Turkic peoples, often leading to the partial or complete loss of their native languages. At this time, the use of Arabic was limited to religious services in mosques, while in other social spheres, Turkish was strictly used. This period lasted until World War I. Thus, dialectal Arabic borrowed many Turkish words, particularly evident in the Damascus dialect, since the Damascus Vilayet became one of the first Ottoman provinces, and was later named one of the main administrative-territorial vilayets. During World War I, the territory of Syria was liberated from the Ottoman Empire. From 1918 to 1920, it existed under the rule of King Faisal I. In 1920, Syria was conquered by France. As part of its colonial policy, France introduced French as a second official language alongside Arabic. French was used in official business, education, and the media. At the lexical level, the colloquial version of the language in Syria began to be enriched with French borrowings. Many words and expressions in the administrative, educational and cultural spheres were borrowed from the French language. The introduction of the French education system increased the number of French speakers in Syria. In 1946, Syria gained independence, and the influence of the French language began to decline. However, many French words, primarily borrowed orally, remained in the language due to the long-lasting impact of the colonial period. French is still taught in Syrian schools.

In the modern world, English holds significant importance as an international language. A large number of Anglicisms exist in all languages of the world, and in the case of the Arabic language, these lexemes can be considered pan-Arabic, since they are used in all dialects of the Arabic language.

Several dialects with unique pronunciation features, specific vocabulary, and distinct syntactic structures have developed in Syria. These include the Aleppo dialect, the coastal dialects of Latakia and Tartus, the Damascus dialect, the Homs dialect, and others. A review of modern scientific works on the Syrian dialect of Arabic reveals that many studies focus on identifying general phonetic patterns and differences between the Syrian dialect and literary Arabic, for example, the exchange of sounds: the replacement of the sound ق [k] with a glottal stop is found in some cities Syria and Lebanon; sound transposition: changing the order of two sounds within a single word. For example, جبذ [jabaz] instead of جذب [jazab] 'pull'; elision: dropping one or more sounds in a word, for example, dropping a glottal stop: مرا [mara] instead of امرأة [mra'a] 'woman'; adding a sound to the original structure of a word: بع [b''] 'sell' (imperative form) — بيع [bi']; changing the composition of the vowels of a lexical unit: مُحَمَّد [muhammad] becomes مُحَمَّد

[mhammad] ‘Muhammad’ [6]. At the grammatical level [7–10] such tendencies are noted as the formation of a new word by combining two or more literary bases (composition): أَيْش [aesh] ‘what’ comes from أَيَّ شَيْءٍ [aj] шaj’ъ ‘any thing’; frequent use of obsolete models: the active participle formed from verbs consisting of more than three letters in the dialect usually begins with a consonant, which corresponds to the model of the Aramaic language, in contrast to literary Arabic, which requires the use of hamza (همزة) as an auxiliary letter; the emergence of new derivational models: تَفَعَّلَ [tafaalal] — تَمَلَّعَ [tmalaan] ‘cheated’; changes in the verb system and expression of the category of time. Thus, E. de Blasio in his work devoted to the analysis of the verb system of the dialect of the city of Tartus in Syria [7] examines a number of dialectal features presented in the following conclusions: 1) negative constructions are most often formed using the prefix ما [ma] in combination with pronoun suffixes: ما شرب [ma shereb] ‘did not drink’; 2) in the imperative form, a reduction of the vowel is observed, which distinguishes this dialect from Damascus: امشي [imshi] — مشي [mshi] ‘go’; 3) the present tense is indicated by the prefix عم, which is sometimes accompanied by the prefix ب [b’]: عم يبيلعب [am bilab] ‘plays’; 4) the future tense is indicated by the prefixes رح [rah] and ح [h]: مارح يشرب [ma rah eshrab] ‘will not drink’. We have noted a number of works devoted to a detailed analysis of individual lexical units that can be useful from a practical point of view in language teaching; paremiology, lexicography and translation, but in a broad sense they are limited - they do not describe the structure of the dialect as a whole. We have also noted discursive studies of the Syrian dialect of Arabic [11–16].

Comparative Studies of Arabic Dialects

The linguistic situation in the Arab world is complex and requires careful attention in linguistic research. Inherent diglossia, along with established sociocultural norms, political processes, and globalization, creates a complex linguistic system. Given this complexity, sociolinguistic analysis is a common approach to studying the linguistic landscape of Arab countries.

The following works, along with numerous studies on the linguistic features of individual Arabic dialects, highlight the fragmentation and lack of systematicity in Arabic studies. In comparative studies of Arabic dialects, the following aspects are frequently encountered:

1. Identification of common phonetic patterns and differences from literary Arabic [17];
2. Study of morphological and syntactic changes in dialects [18; 19];
3. Examination of the lexical level, including borrowings [20–23];
4. Discursive studies [24; 25].

A significant work analyzing the language situation in Arab countries is Reem Bassiouney’s “Arabic Sociolinguistics” [2]. This work discusses diglossia, analyzes the distribution of functions between the two forms of Arabic in public life, and examines the social and psychological factors influencing speakers’ choices of

language forms. It also examines the impact of gender differences on Arabic usage, analyzes speech behaviors and strategies characteristic of each gender, and considers sociocultural norms and stereotypes related to language and gender. The paper describes Arabic's role in forming individual and group identity and analyzes how linguistic variation affects the expression of ethnic, religious, and national identities. The author examines language policy in Arab countries and its impact on the status of Arabic, as well as its use in media, education, and public administration.

For instance, in 2008, Mazin al-Mubarak, a scholar and member of the Academy of the Arabic Language in Damascus, advocated for strengthening the status of Modern Standard Arabic (MSA) and gradually eradicating the use of dialects in colloquial speech in Syria. His plan included information programs for families, increased use of MSA in media, more songs, plays, and films in MSA, and a rejection of works not presented in literary Arabic, even if they were entertaining.

A key problem in analyzing Arabic dialects is the difficulty in obtaining linguistic material. Unlike literary Arabic, which has extensive databases of texts, grammars, and dictionaries, dialects are primarily oral, complicating data collection and analysis. To address this issue, researchers often rely on: 1) Secondary data (linguistic material collected by other researchers); 2) A lack of source indications for linguistic material.

By illustrating the linguistic situation in Arab countries through examples from Egyptian, Levantine, Maghreb, Mesopotamian, Sudanese, Yemeni, Hejazi, and Maltese dialects, the author draws several key conclusions [27]:

1. Plurality of dialects: Arabic is characterized by linguistic diversity, with many dialects varying by region, social group, and other sociocultural factors.
2. Phonetic variations: Different dialects show variations in phoneme pronunciation, leading to difficulties in speech recognition among speakers of different dialects.
3. Morphological differences: Dialects vary in their use of prefixes, suffixes, and clitics.
4. Syntactic differences: Dialects vary in word order usage.
5. Lexical and semantic differences: Dialects feature unique borrowings from other languages and many lexical units from literary Arabic that carry different meanings in dialects, along with changes in the phonetic structure of lexemes.

This work is based on linguistic material from monographs, scientific articles, fiction, and journalistic texts, aiming to: 1) Define the concept of dialect; 2) Describe and illustrate dialectal features across different Arabic dialects.

The author argues that literary Arabic functions in official contexts but is not the native language of its speakers. Dialectal Arabic is defined as an informal variant used in daily life, characterized by distinctive features at all language levels. This statement is controversial, as it addresses fundamental issues of linguistic identity and the formation and functioning of language in society. Such a claim

requires careful justification and evidence, as it challenges traditional views on the role and status of literary Arabic.

P.E. Demina's study [25] examines the features of international communication and professional foreign language communication across four dialects: Syrian, Omani, Egyptian, and Iraqi. The author identifies difficulties in interacting with native Arabic speakers and suggests ways to overcome these challenges. Key problems include a lack of dialect knowledge, the principle of speech economy observed at all levels, non-compliance with literary Arabic rules, pronunciation changes due to dialect phonetic systems, and the use of non-standard abbreviations.

The author's analysis of the language situation and empirical data on four Arabic dialects leads to the following conclusions:

1. Diglossia affects language functioning and interpretation in various communication contexts;

2. Each Arabic dialect has distinct phonetic, lexical, and morphological features that differentiate it from literary Arabic;

3. The choice of language variant depends on the communication context: literary Arabic is used in official meetings, while dialects dominate everyday conversations;

4. Vocabulary heterogeneity: Dialects feature borrowed vocabulary and many lexical units from literary Arabic that carry different meanings in dialects, along with changes in the phonetic structure of lexemes.

The work's debatable nature arises from several factors: most language examples lack source references, and some examples are drawn from other studies (as noted by the author).

Yasin Demir's study [13] provides a general description of the Syrian dialect of Arabic, highlighting its specific usage. The author notes that the vast geographical distribution of Arabic and the rich historical events in the regions where its speakers live create a continuum of dialects that reflect various cultural characteristics. The Syrian dialect, one of the most significant and widely used, plays a crucial role in communication among Syrians, particularly regarding the social integration of Syrian refugees in Europe.

O.I. Redkin's work offers a brief historical overview of the stages in the development of the Arabic language. The author distinguishes between two language variants: classical and modern literary Arabic, emphasizing the concept of diglossia—the coexistence of two language forms: formal (*fuskha*), used in official contexts, and informal (*ammiya*), used in daily communication [26]. This division creates challenges in teaching and learning Arabic.

The article examines the geographical distribution of the Syrian dialect, emphasizing its influences from various languages, including Arabic, Syriac (Aramaic), Greek, Turkish, and French, as reflected in its vocabulary and grammatical structure. The main section analyzes specific communicative situations and common expressions used in everyday life within the Syrian dialect.

The author offers numerous examples of phrases for greetings, expressing agreement or disagreement, farewells, weather descriptions, and other common everyday situations. Each example in the Syrian dialect is compared with its literary Arabic counterpart, aiding in identifying differences and features of dialectal speech.

The study presents language examples, but the selection method remains unclear. The main section consists of a list of expressions and their approximate translations. However, there is no analysis of the structure, semantics, or pragmatics of the presented lexical units.

The modern world is experiencing significant changes due to globalization, technological advances, and migration, all of which impact the linguistic situation and the use and perception of Arabic dialects. The sociolinguistic situation in the Arab world, particularly in Syria, is unique, as multiple varieties of Arabic coexist with English, French, and other languages. In light of the rapidly changing geopolitical situation, continuous updates on the linguistic landscape in Syria are essential. Migration flows, political changes, and the influence of foreign languages directly impact linguistic practices, the spread of dialects, and the functioning of literary Arabic.

Electronic tools for recognizing and transforming dialects into a literary version of the language

In the development of a multimedia dictionary for the Syrian dialect of Arabic, various studies on creating online resources for dialect “normalization” and electronic dictionaries were analyzed. Researchers are utilizing transformer-based models, specifically BERT and GPT-2, to address and overcome the linguistic differences between Arabic dialects and Modern Standard Arabic (MSA). These differences manifest not only at the phonetic level but also in morphology, grammar, and vocabulary. The MADAR³ corpus, which includes sentences in English and various Arabic dialects, serves as the primary data source for training the model. While results indicate challenges with automatic text normalization, manual analysis reveals that the model performs effectively overall, despite some shortcomings in the quality of the source data.

The Nâbra project has compiled a corpus of Syrian dialects that encompasses 10 territorial dialects and features approximately lexemes sourced from diverse materials, including social media, blogs, films, and local poetry. The primary objective of this project is to prepare a comprehensive corpus to facilitate further linguistic research rather than to create a traditional dictionary. The annotation of the texts was carried out using the Tawseem web interface, which streamlined the process and ensured high-quality annotations, achieving an average F1-score of about 0.85.

³ Madar. URL: <https://sites.google.com/nyu.edu/madar/> (accessed: 21.04.2025).

The Qabas⁴ project offers an open-access resource characterized by a heterogeneous vocabulary that includes both modern literary Arabic and dialectal lemmas. Despite its extensive vocabulary coverage and integration with various linguistic resources, Qabas is noted to lack sufficient etymological information on the lexemes it contains. Furthermore, the resource is primarily designed for native Arabic speakers, as the English translations provided are limited to word equivalents rather than comprehensive definitions.

The ongoing development of electronic dialect resources, such as text corpora, lexicographic databases, and grammatical descriptions, represents a promising direction in contemporary linguistics. These resources enable researchers to conduct empirical analyses of dialect systems, identify statistically significant patterns in their usage, and monitor the dynamics of language change over time. The incorporation of artificial intelligence (AI) technologies into the development and application of these electronic dialect resources significantly enhances researchers' capabilities in addressing practical challenges. Methods such as machine learning and natural language processing facilitate the automation of dialect text analysis, recognition and synthesis of dialect speech, and the creation of tools that support the learning of dialects.

Dictionary of the Syrian Dialect of Arabic

Currently, there is both a need and the technical capability to develop a multimedia dictionary that presents dialect lexemes distinct from literary Arabic, including the nuances of word usage in various contexts. The demand for a system of dialect dictionaries addresses both linguistic and extralinguistic needs, providing access to valuable information and enhancing the quality of research in the field of Arabic language studies.

The integration of audiovisual materials and interactive elements will make the dictionary more user-friendly, facilitating a deeper exploration of dialects and their unique features. The principles of vocabulary selection are outlined in Table 2. Additionally, the option to download data tailored for different categories of users enables the dictionary to serve educational, scientific, and translation purposes effectively.

Furthermore, the model of the multimedia dialect dictionary and the methodology for analyzing lexical units can be adapted for the lexicography of other Arabic dialects.

The dictionary implements a number of functions (Table 3).

⁴ Birzeit University. Qabas. URL: <https://sina.birzeit.edu/qabas/> (accessed: 21.04.2025).

Table 2. Principles of vocabulary selection

Principle	Implementation
The Principle of Dialectal Affiliation	Inclusion of lexemes that are unique to a specific dialect or group of dialects, which are not found in the literary language.
The Principle of Semantic Significance	Semantically significant vocabulary forms the basis of the linguistic picture of the world and reflects key cultural concepts. This includes words denoting the main spheres of human activity: work, communication, basic emotions, and objects of both material and non-material culture.
The Principle of Thematic Relevance	Inclusion of words related to specific topics or areas of knowledge, facilitating the formation of lexical-semantic groups.
The Principle of Chronological Relevance	Accounting for changes in vocabulary over a certain period, reflecting the evolution of language.
The Principle of Territorial Limitation	Consideration of the geographical distribution of words, including lexical units used within specific regions or localities, and recording information about the area of use in dictionary entries.
The Principle of Taking into Account Sociolinguistic Factors	Recording information about the social characteristics of dialect speakers (age, education, occupation, etc.), particularly relevant when conducting surveys to collect linguistic material.
The Principle of Reflecting Options	Documentation of different phonetic, morphological, and semantic variants of the same word across various parts of the dialect area.
The Principle of Ethnographic Significance	Inclusion of lexical units that denote unique elements of the material and spiritual culture of a region (e.g., household items, rituals, customs, folk beliefs), with a presumption of incorporating multimedia elements.
The Principle of Source Analysis	Utilizing a variety of sources to identify and select lexical units for inclusion in the dictionary.
The Principle of Audience Orientation	Considering the level of training, interests, and language skills of users (e.g., students, researchers, native speakers) to adapt the presentation of material to meet their needs.

Source: compiled by Aida D. Haddad & Natalia V. Novospasskaya.

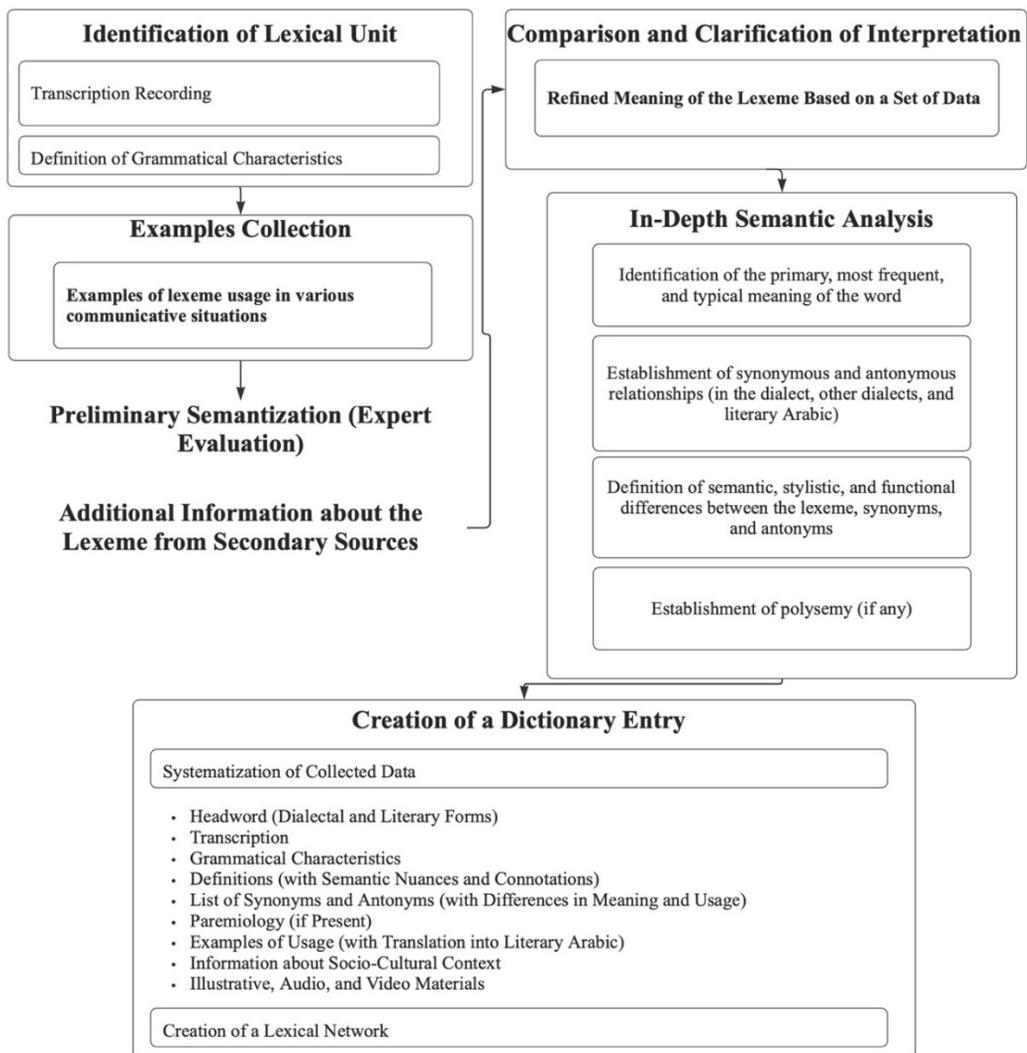
Table 3. Dictionary functions

Function	Description	Implementation
Informative Function	Providing the user with information about a word: interpretation, transcription, grammatical commentary, etymology, synonyms, antonyms, context of use	Text definitions, examples of usage in context, transcription
Attractive Function	Capturing the user's attention and maintaining their interest	Use of multimedia content (audio, video, images)
Educational Function	Assisting the user in learning new lexical units and expanding their vocabulary	Interactive exercises, tests, flashcards, examples of usage in context, audio and video materials for practicing pronunciation, and the ability to download based on formulated requests

End of Table 3

Function	Description	Implementation
Cultural Memory Preservation Function	Preserving and transmitting information about cultural heritage, traditions, customs, and values associated with the dialect	Inclusion of information about cultural contexts of word usage, folklore texts, rituals, and traditions in dictionary entries, as well as visualization of cultural phenomena related to the dialect for clearer representation of information
Updating and Documenting Function	Recording the current state of the dialect, reflecting new words and expressions emerging in the speech of native speakers	Ability to edit existing dictionary entries and create new ones

Source: compiled by Aida D. Haddad & Natalia V. Novospasskaya.



The process of semantization of a lexeme and compilation of a dictionary entry

Source: compiled by Aida D. Haddad & Natalia V. Novospasskaya.

Conclusions

The multimedia dictionary of the Syrian dialect aims to address current challenges in the systematization and lexicography of Arabic dialectal vocabulary. It is a valuable tool for both practical study and research in the Arabic language, contributing to the preservation and popularization of the dialect. Additionally, the dictionary gives opportunities for cross-cultural research. By comparing the Syrian dialect with other Arabic dialects and languages, as well as examining its features within various social and cultural contexts, researchers can establish cultural and historical connections between different regions.

References / Библиографический список

1. Tuleubaeva, S. (2019). On the Modern Language Situation in Arabic Countries. *Eurasian Arabic Studies*, 2(8), 45–56. (In Russ.). EDN: RUVQZQ
Тулеубаева С.А. О современной языковой ситуации в арабских странах // Арабистика Евразии. 2019. Т. 2. № 8. С. 45–56. EDN: RUVQZQ
2. Bassiouney, R. (2020). *Arabic Sociolinguistics: Topics in Diglossia, Gender, Identity, and Politics*. Washington: George town University Press. <https://doi.org/10.2307/j.ctv10kmbxg>
3. Dubinina, N.V., Merai, M.I., & Borisenko, E.A. (2024). Characteristics of Interethnic Communication of Native Speakers of Arabic. *Cognitive Studies of Language*, 1–1(57), 114–117. (In Russ.). EDN: UEEQTF
Дубинина Н.В., Мерай М.И., Борисенко Е.А. «Белый диалект» как язык межнационального общения в арабском языковом пространстве // Когнитивные исследования языка. 2024. № 1–1 (57). С. 114–117. EDN: UEEQTF
4. Zarytovskaya, V.N., & Al-Rahbi, A.M. Omani Arabic: Historical and Linguistic Description. *Philology. Theory & Practice*, 12(10), 302–307. (In Russ.). <https://doi.org/10.30853/filnauki.2019.10.64> EDN: FJLUBB
Зарытовская В.Н., Аль-Рахби А.М. Оманский диалект арабского языка: историко-лингвистическое описание // Филологические науки. Вопросы теории и практики. 2019. Т. 12. № 10. С. 302–307. <https://doi.org/10.30853/filnauki.2019.10.64> EDN: FJLUBB
5. Chagal, V.E. (1998). *Arab Countries: Language and Society*. Moscow. (In Russ.).
Шагаль В.Э. Арабские страны: язык и общество. М., 1998.
6. Khrisat, A., Alharthy, Z. (2015) Arabic Dialects and Classical Arabic Language. *Advances in Social Sciences Research Journal*, 2(3), 254–260. <https://doi.org/10.14738/assrj.24.1048>
7. Blasio, E. (2022). A Preliminary Study on the Verbal System of the Syrian Dialect of Ṭartūs. *Annali di Ca' Foscari. Serie orientale*, 58, 183–198. <https://doi.org/10.30687/AnnOr/2385-3042/2022/01/007>
8. Bernikova, O., & Redkin, O. (2019). Digital Humanities: The Case Study of Arabic Dialectology. In: *IMSCI 2019 — 13th International Multi-Conference on Society, Cybernetics and Informatics, Proceedings: 13, Orlando, FL, 06–09.07.2019* (vol. 1, pp. 1–60). Orlando. EDN: MGOAHI
9. Faraj, R. (2020). Morphosyntactic Transfer from a Spoken Heritage Dialect to the Standard Variety: The Case of Heritage Speakers of Syrian Arabic. *Heritage Language Journal*, 17, 202–233. <https://doi.org/10.46538/hlj.17.2.5> EDN: JVVLLN
10. Nayouf, A., Hammouda, T., Jarrar, M., Zaraket, F., & Kurdy, M.B. (2023). Nâbra: Syrian Arabic dialects with morphological annotations. *arXiv*, Article 2310.17315. <https://doi.org/10.18653/v1/2023.arabicnlp-1.2>

11. Ramos, C. (2022). Idioms in Syrian Arabic: a Semantic and Grammatical Approach to the Verb. *Yearbook of Phraseology*, 13, 161–188. <https://doi.org/10.1515/phras-2022-0009> EDN: ZSMZXT

12. Hodeib, C. (2023). Conceptualizations and Evaluations of (Im)Politeness in Syrian Arabic. *Journal of Politeness Research*, 20. <https://doi.org/10.1515/pr-2020-0016> EDN: PJYQ GK

13. Demir, Y. (2022). General Characteristics of the Syrian Arabic Dialect and Examples of a Common Usage in Daily Life. *European Journal of Applied Linguistics Studies*, 5. <https://doi.org/10.46827/ejals.v5i1.340> EDN: TKBSBY

14. Haddad, A.D., & Khalil, A.A.A. (2025). Lexicography of the Syrian Dialect of Arabic: Sources and Dictionary. *Training, Language and Culture*, 9(2), 20–30. <https://doi.org/10.22363/2521-442X-2025-9-2-20-30> EDN: LDJCVS

15. Khalil, A.A.A., & Larina, T.V. (2022). Terms of Endearment in American English and Syrian Arabic Family Discourse. *RUDN Journal of Language Studies, Semiotics and Semantics*, 13(1), 27–44. <https://doi.org/10.22363/2313-2299-2022-13-1-27-44> EDN: RUITPY

16. Novospasskaya, N.V., & Haddad, A.D. (2024). Film Discourse: a Resource for Creating a Dialect Dictionary. *Philology and Culture*, 4(78), 139–145. <https://doi.org/10.26907/2782-4756-2024-78-4-139-145> EDN: CWKELR

17. Kulikov, V., & Hazi, R. (2024). Phonetic Realization of the Emphasis Contrast in Voiceless Sops by Speakers of Four Arabic Varieties: The Case of F2 and VOT. *Arabic Linguistics*, 8, 1—31. <https://doi.org/10.1075/arli.00001.kul> EDN: JXOJNZ

18. Krasnoschekov, E.V. (2021). Teaching ESP for Law Students at Universities in Russia and Abroad. *Language and Culture*, 55, 103–114. (In Russ.). <https://doi.org/10.17223/19996195/55/7> EDN: ZZOGVS

Краснощеков Е.В. Конструкции с идафой как средство выражения притяжательности в диалектах арабских языков // Язык и культура. 2021. № 55. С. 103–114. <https://doi.org/10.17223/19996195/55/7> EDN: ZZOGVS

19. Hallman, P., & Al-Balushi, R. (2022). Pronominalization and Clitic Doubling in Syrian and Omani Arabic. *Linguistics*, 60(5), 1295–1336. <https://doi.org/10.1515/ling-2019-0053>

20. Kalinina, G.S. (2021). Features of Clichéd Vocabulary in Classical Arabic and Its Dialects (in the Intentions “Greetings”, “Gratitude”, “Advice”, “Invitations”, “Congratulatory”, “Apologies” and “Farewells”). *Philology and Culture*, 1(63), 46–53. (In Russ.). <https://doi.org/10.26907/2074-0239-2021-63-1-46-53> EDN: VLDZAN

Калинина Г.С. Особенности клишированной лексики в классическом арабском языке и его диалектах (в интенциях «приветствие», «благодарность», «совет», «приглашение», «поздравление», «извинение», «прощание») // Филология и культура. 2021. № 1 (63). С. 46–53. <https://doi.org/10.26907/2074-0239-2021-63-1-46-53> EDN: VLDZAN

21. Antonian, A.S. (2024). Loanwords in the Arabic Dialects: Eponymisation Aspect. *Eurasian Arabic Studies*, 1, 16–30. (In Russ.). <https://doi.org/10.26907/2619-1261.2024.7.1.16-30> EDN: UTKYXU

Антонян А.С. Заимствования в арабских диалектах: аспект эпонимизации // Арабистика Евразии. 2024. № 1. С. 16–30. <https://doi.org/10.26907/2619-1261.2024.7.1.16-30> EDN: UTKYXU

22. Cherkasova, A.P. (2012). Innovations in Arabic Dialects of Maghreb Countries Caused by Interaction of the French and Arabic Languages. *Science Journal of Volgograd State University. Linguistics*, 1, 242–245. (In Russ.). EDN PCAFOX

Черкасова А.П. Инновации в арабских диалектах стран Магриба как результат взаимодействия французского и арабского языков // Вестник Волгоградского государственного университета. Серия 2: Языкознание. 2012. № 1. С. 242–245. EDN: PCAFOX

23. Antonian, A.S. (2024). Orthographic and Morphological Variation of Loanwords in the Arabic Vernaculars. *Vestnik of Saint Petersburg University. Asian and African Studies*, 16(1), 81–99. (In Russ.). <https://doi.org/10.21638/spbu13.2024.106> EDN: IMLSYI

Антоян А.С. Орфографическая и морфологическая вариативность заимствований в разговорных диалектах арабского языка // Вестник Санкт-Петербургского университета. Востоковедение и африканистика. 2024. Т. 16. № 1. С. 81–99. <https://doi.org/10.21638/spbu13.2024.106> EDN: IMLS YI

24. Dugalich, N.M., Shavtikova, A.T., & Izildin, O. (2023). Strategies for Creating the Image of a Politician in an Arabic Polycode Text of the Series “اختيار”. *RUDN Journal of Language Studies, Semiotics and Semantics*, 14(3), 946–959. <https://doi.org/10.22363/2313-2299-2023-14-3-946-959> EDN: ISLZAW

25. Demin, P.E. (2020). Specificities of International and Foreign Language Professional Communication in the Face of a Diversity of Arabic Dialects. *Political Linguistics*, 5(83), 149–156. (In Russ.). <https://doi.org/10.26170/pl20-05-13> EDN: KBGKYC

Демин П.Е. Особенности международного общения и профессиональной иноязычной коммуникации в условиях разнообразия диалектов арабского языка // Политическая лингвистика. 2020. № 5 (83). С. 149–156. <https://doi.org/10.26170/pl20-05-13> EDN: KBGKYC

26. Redkin, O., & Bernikova, O. (2022) Key Milestones in the History of the Arabic language: the Dialectics of Development. *Eurasian Arabic Studies*, 5(4), 16–31. (In Russ.). <https://doi.org/10.26907/2619-1261.2022.5.4.16-31> EDN: NZDTZF

Редькин О.И., Берникова О.А. Основные вехи в истории арабского языка: диалектика развития // Арабистика Евразии. 2022. № 5 (4). С. 1–31. <https://doi.org/10.26907/2619-1261.2022.5.4.16-31>

27. Alnosaire, A., & Sartini, N. (2021). A Sociolinguistics Study in Arabic Dialects. *PRASASTI: Journal of Linguistics*, 6(1). <https://doi.org/10.20961/prasasti.v6i1.43127>

Information about the authors:

Natalia V. Novospasskaya, PhD in Philology, Associated Professor, Associated Professor at the General and Russian Linguistics Department, Faculty of Philology, RUDN University (6 Miklukho-Maklaya Str., Moscow, Russian Federation, 117198); *Research interests*: terminology, lexicography, typological linguistics; *e-mail*: novospasskaya_nv@pfur.ru
ORCID: 0000-0001-7599-0246; SPIN-code: 1362-1820, AuthorID: 397500; Scopus Author ID: 57219297279; ResearcherID: 8769-2019.

Aida D. Haddad, PhD student in Philology, assistant at the General and Russian Linguistics Department, Faculty of Philology, RUDN University (6 Miklukho-Maklaya Str., Moscow, Russian Federation, 117198); *Research interests*: dialectology, Arabic language, lexicography, typological linguistics; *e-mail*: khaddad_a@pfur.ru ORCID: 0000-0002-7765-2133; SPIN-code: 1601-8469, AuthorID: 1156296.

Сведения об авторах:

Новоспасская Наталья Викторовна, кандидат филологических наук, доцент, доцент кафедры общего и русского языкознания, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: терминология, лексикография, типологическое языкознание; *e-mail*: novospasskaya_nv@pfur.ru
ORCID: 0000-0001-7599-0246; SPIN-код: 1362-1820, AuthorID: 397500; Scopus Author ID: 57219297279; ResearcherID: 8769-2019.

Хаддад Аида Джихадовна, кандидат филологических наук, ассистент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: диалектология, арабский язык, лексикография, типологическое языкознание; *e-mail*: khaddad_a@pfur.ru
ORCID: 0009-0009-5295-0802; SPIN-код: 1601-8469, AuthorID: 1156296.