

The Online Atlas of the Languages of Iran: Design, Methodology and Initial Results

Vol. 12, No. 2, Tome 62
pp. 231-291
June & July 2021

Received: 25 June 2019
Received in revised form: 4 December 2019
Accepted: 7 December 2019

Author name	Academic rank, field / Department	University
1. Mortaza Taheri-Ardali	Assistant Professor in General Linguistics, Department of English; Bakhtiari Studies Research Institute	Shahrekord University, Shahrekord, Iran
2. Erik Anonby	Professor in Linguistics and French	Carleton University, Ottawa, Canada
3. Amos Hayes	Technical Manager, Geomatics and Cartographic Research Centre (GCRC)	Carleton University, Ottawa, Canada
4. Zahra Azin	PhD candidate in Linguistics, Language Documentation and Revitalization	Carleton University, Ottawa, Canada
5. Zhino Ebrahimi	Master Student in Cognitive Science, Departement of Social Sciences	University of Kaiserslautern, Kaiserslautern, Germany
6. Ergin Öpengin	Lecturer	University of Kurdistan-Hewlêr, KRI, Iraq
7. Noush Azarin Atabaki	MA (graduate) in General Linguistics	University of Tehran, Iran
8. Adam Stone	Research Associate, Geomatics and Cartographic Research Centre (GCRC)	Carleton University, Ottawa, Canada
9. Donald Stilo	Scientist (retired)	Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
10. Elmira Esmaeelpour	PhD candidate in General Linguistics	Alzahra University, Tehran, Iran
11. Maryam Amani-Babadi	PhD (graduate) in General Linguistics	Payame Noor University, Tehran, Iran
12. Muhammed Ourang	PhD candidate in Linguistics	The University of New South Wales (UNSW), Sydney, Australia
13. Elham Izadi	PhD (graduate) in General Linguistics	Bu-Ali Sina University, Hamedan, Iran

Language Related Research *12(2), (June & July, 2021) 231-291*

Author name	Academic rank, field / Department	University
14. Robert Oikle	Geomatics Researcher, Geomatics and Cartographic Research Centre (GCRC)	Carleton University, Ottawa, Canada
15. Nawal Bahrani	MA (graduate) in General Linguistics	Allameh Tabataba'i University, Tehran, Iran
16. Habib Borjian	Research Scholar	Columbia University, New York, US
17. Ava Bahrami	MA (graduate) in TEFL	Sharif University of Technology, Tehran, Iran
18. Hamideh Poshtvan	MA (graduate) in General Linguistics	Alzahra University, Tehran, Iran
19. Shiva Piryaee	PhD candidate in Linguistics	Tarbiat Modares University, Tehran, Iran
20. Peiman Pishyardehkordi	PhD student in Linguistics, Department of Linguistics and English Language	University of Canterbury, Christchurch, New Zealand
21. Parisa Sabethematabadi	PhD candidate in Applied Linguistics;	Carleton University, Ottawa, Canada
	Instructor of English for Applied Purposes	Algonquin College, Ottawa, Canada
22. Shima Jaafari-Dehaghi	Assistant Professor in Ancient Iranian Languages, Persian Language and Literature Department	Velayat University, Iranshahr, Iran
23. Maryam Jalili Doab	MA in (graduate) General Linguistics	Semnan University, Semnan, Iran
24. Faranak Jamaledin	PhD candidate in Linguistics	Ferdowsi University of Mashhad, Iran
25. Kamyar Joulaei	PhD (graduate) in General Linguistics	Institute for Humanities and Cultural Studies, Tehran, Iran
26. Hasan Haji Shabaniyan	MA (graduate) in General Linguistics	Islamic Azad University of Central Tehran Branch, Tehran, Iran
27. Jila Khanjani	Assistant Professor in General Linguistics	Payame Noor University, Department of Literature and Foreign Languages, P.O. box: 193953697, Tehran, Iran
28. Leila Dianat	PhD (graduate) in General Linguistics	Razi University, Kermanshah, Iran
29. Ali Rashidi	MA (graduate) in Linguistics	Sistan va Baluchestan University, Zahedan, Iran

Author name	Academic rank, field / Department	University
30. Reza Rezvani Borujeni	MA (graduate) in Linguistics	Islamic Azad University of Khomein, Iran
31. Zohreh Rahnema	MA (graduate) in Persian Language and Literature	Islamic Azad University of Zanjan, Iran
32. Zeinab Zamani-Gandomani	MA student in TEFL	Shahrekord University, Shahrekord, Iran
33. Laurentia Schreiber	Research Fellow in General Linguistics	University of Bamberg, Bamberg, Germany
34. Negar Sherafat	MA (graduate) in Linguistics	Allameh Tabataba'i University, Tehran, Iran
35. Jaffer Sheyholislami	Professor in Applied Linguistics and Discourse Studies; cross-appointed in Linguistics, and Cultural Mediation	Carleton University, Ottawa, Canada
36. Mina Salehi	MA (graduate) in Ancient Iranian Languages and Culture	Allameh Tabataba'i University, Tehran, Iran
37. Mahnaz Talebi-Dastenaee	PhD (graduate) in General Linguistics	Alzahra University, Tehran, Iran
38. Ali Akbar Abedian Kaspary	MA (graduate) in Linguistics	Managing Editor of Publishing House, Tehran, Iran
39. Maryam Sadat Ghiasian	Associate Professor in General Linguistics	Payame Noor University, Tehran, Iran
40. Mehdi Fattahi	Assistant Professor in General Linguistics	Razi University, Kermanshah, Iran
41. Saeideh Ghandi	PhD candidate in General Linguistics	Tarbiat Modares University, Tehran, Iran
42. Mojtaba Gheitasi-Doolabi	PhD candidate in TEFL	Ilam University, Ilam, Iran
43. Roya Kamali	MA (graduate) in General Linguistics	High School English Teacher, Tehran, Iran
44. Farzaneh Goshtasb	Associate Professor in Ancient Iranian Languages and Culture	Institute for Humanities and Cultural Studies, Faculty of Linguistics, Department of Ancient Iranian Languages and Culture, Tehran, Iran
45. Hassan Mohebbi Bahmani	MA (graduate) in General Linguistics	Imam Khamenei Technical Faculty, Minab, Iran
46. Mohammad Mohammadi	MA (graduate) in General Linguistics	Iranian Ministry of Education
47. Ronak Moradi	Assistant Professor in General Linguistics, Department of Linguistics and Foreign Languages	Payame Noor University, Tehran, Iran

Author name	Academic rank, field / Department	University
48. Mehrdad Meshkinfam	PhD (graduate) in General Linguistics	Bu-Ali Sina University, Hamedan, Iran
49. Razieh Mousavi Khoo	PhD candidate in Ancient Iranian Languages and Culture	Azad University of Tehran - Science and Research Branch, Iran
50. Arefe Nezhad Bahram	MA (graduate) in General Linguistics	Allameh Tabataba'i University, Tehran, Iran
51. Fatemeh Nemati	Assistant Professor in General Linguistics	Persian Gulf University, Bushehr, Iran
52. Maryam Nourzaei	Postdoctoral Researcher	Uppsala University, Uppsala, Sweden and Otto-Friedrich-Universität Bamberg, Bamberg, Germany
53. Emily Wang	MA (graduate) in Education	Carleton University, Ottawa, Canada
54. Hossein Hashemi-Zarajabad	PhD candidate in General Linguistics	Allameh Tabataba'i University, Tehran, Iran

Abstract

Iran is home to a treasury of languages representing diverse language families: Iranian, Turkic, Semitic, Indic, Dravidian, Armenian, and Kartvelian, as well as sign languages. Despite valuable research carried out by Iranian and western scholars, there is still no comprehensive publication depicting the geographic distribution and linguistic status of language varieties in Iran. In order to work toward this goal, the *Atlas of the Languages of Iran* (ALI) (www.iranatlas.net) was officially launched in 2015 as an international, online research programme. The present study opens with a historical overview of the research context and underlines the ongoing necessity of constructing such an atlas for the languages of Iran. The article then outlines the design of the Atlas, the methodology necessary to guide and organize the research project, and the results which are being generated. Specifically, we look at the Atlas bibliography, classification of Iran's languages, the design of a set of linguistic data questionnaires, and findings from fieldwork. Selected initial research results from fieldwork in Chahar Mahal va Bakhtiari and Hormozgan Provinces, where the research process is most advanced, are provided in the form of linguistic data maps.

Keywords: Languages of Iran, Language documentation, Language atlas, Language map, Language classification

1. Introduction

Over the past decades, many efforts have been made to explore and understand the linguistic situation in Iran. However, no systematic study has been made available to the public on the basis of which a coherent picture of the country's languages can be drawn. In order to provide a more comprehensive account of the language situation as well as to document the linguistic varieties in Iran, a project called the *Atlas of the Languages of Iran* (ALI) was initiated in 2009. Initially based at Carleton University, it has become an international research programme with a majority of colleagues and contributors based in Iran. The online version of the atlas launched in 2015 at www.iranatlas.net and is continually updated as new sections of the atlas are completed.

2. ALI and its ultimate goal

The *Atlas of the Languages of Iran*, or ALI, is the title of an online research programme that documents language varieties within the geographical boundaries of Iran and prepares maps of language distribution and linguistic data from this area. In ALI, researchers and consultants help advance the atlas. At various stages, about 90 researchers have contributed to the atlas as volunteers, most often in the form of provincial teams: currently 19 of the country's 31 provinces are active. ALI uses the Nunaliit platform as the basis for its online interface. This open source platform is freely available to the public and can be used to prepare any linguistic atlas. Nunaliit's relational structure is flexible and can help overcome common challenges and gaps in the construction of an atlas.

The ultimate goal of the ALI research programme is to provide an overview of the country's language situation. To work toward this end, project members are working on language distribution maps, linguistic data maps, and preparation of a corpus consisting of audio and video files from

language varieties across Iran. We also aim to answer the following questions:

- What are the salient linguistic features of Iran's languages and dialects, and how are these linguistic features distributed geographically?
- In which parts of Iran are these language varieties spoken, and how does this compare to the distribution of linguistic features?
- How do scholars and linguists classify these languages, and how can current classifications be improved?

3. Linguistic classification and bibliography

One of the first contributions of the ALI programme is the assembly of a comprehensive working classification of languages in Iran, provided to users through the Atlas website. This tool builds on the prior achievements of many researchers in the field of Iranian linguistics. ALI uses two models to depict the classification of languages in Iran. The first model is a traditional classification tree. In this classification, more than 500 distinct language varieties have been catalogued in Iran, of which approximately 400 belong to Iranian (Iranian) language family. The second model is a multi-dimensional language relation web, built using a force-directed graphic visualization. The second model can be used alongside the traditional classification to show connections between language varieties, but it has the added advantage of explicitly recognizing structural similarity and social identification as classificatory devices alongside the output of the comparative method. From the outset of the ALI programme, the atlas website has also featured a continually growing bibliography of scientific works related to linguistic data, language distribution, and maps of Iran's languages.

4. Methodology

The process of construction for language distribution maps proceeds according to region: each province is investigated, analyzed, and finalized separately. To work toward this goal, we first download a list of settlements in each province from the website of the Iranian Statistics Centre, in the form of an Excel file for each provincial sub-district, including all cities and villages. After transcribing the Persian settlement names in roman script, we collect local pronunciations of place names and transcribe them in a simplified phonemic transcription system. At the same time, we conduct research into the language distribution. First, the research team assembles a comprehensive list of language varieties in the province. Linguists and others familiar with the languages of the region, such as teachers, are then asked for an overview of languages spoken in each of the settlements and an estimate of the proportion of each language spoken there as a mother tongue. After entering the language distribution percentage for each settlement along with the place name transcriptions in the Excel file, the prepared data are transferred to the ALI database and associated with the settlement's location on the country-level map provided on the ALI website.

5. Linguistic data collection

The main tool for linguistic data collection in this research programme is the ALI linguistic data questionnaire. Once the language distribution in each province has been established and visualized on a province-level map, it is possible to more efficiently select linguistically and geographically representative research sites for linguistic data collection. The questionnaire, central to the research programme as a whole, has been designed for systematic collection of salient linguistic structures from languages across Iran. A pilot version was designed by Erik Anonby under the valuable guidance of Donald Stilo and other colleagues and carried out in about 50 locations between 2015 to 2018. The questionnaire has since been

significantly updated. Currently, it is organized into four sections: introduction and sociolinguistic context, lexicon, morphosyntactic, and numbers. Topics in phonology and semantics have been integrated into the other sections, and are also addressed by means of oral text collection and analysis.

6. The output of ALI project

Key outputs of the ALI project are language distribution maps at the provincial and the country levels, maps related to linguistic data, and associated media. Audio and video recordings of various genres such as folktales, narratives, and poems from selected language varieties will be uploaded on the website and will be freely available to users. As is important for any language atlas, we have focused on building language distribution maps for all parts of the country. This will eventually allow for a first integrated map of the country's languages. Linguistic data collection has also started in five provinces. Questionnaires have been conducted in Chahar Mahal va Bakhtiari Province (26 locations), Hormozgan Province (10), Ilam Province (5), Zanzan Province (2), and Semnan Province (3). Since a first phase of linguistic data collection from Chahar Mahal va Bakhtiari Province has been completed, we have embarked on analysis and publication of different parts of the questionnaire. For Chahar Mahal va Bakhtiari Province, a total of five maps so far have been extracted from the linguistic analysis, of which three are related to lexical diversity and two introduce elements of phonological diversity in the province.

7. Results

The pilot phase of the ALI research programme has been operational for more than ten years now. We are now entering a full-scale atlas phase and aim to bring this atlas to fulfillment over the next two decades. The time

required will vary depending on the availability and contributions of team members; participation of project partner institutions; funding; and other possible factors. Due to the time-consuming nature of atlas-related tasks, we prioritize the release of findings to the public after the completion of each of its sections, including language distribution data, linguistic data, and bibliographic information.

In this article, by examining and explaining the different parts of the ALI research programme, we have sought to provide a relatively comprehensive picture of how the project is structured and how it is moving forward. A foundational factor in its successes so far is an unwavering commitment to provide the highest possible level of quality in the different dimensions of the work. Research on language distribution and linguistic data collection in remaining provinces, along with a growing bibliography of scholarly resources, are among the key tasks ahead for the *Atlas of the Languages of Iran*.

