

Research Overview

Erik Anonby, Mortaza Taheri-Ardali and Amos Hayes

The *Atlas of the Languages of Iran* (ALI): A Research Overview

There have been a number of important efforts to map out the languages of Iran, but until now no language atlas, or even a comprehensive and detailed country-level language map, has been produced. One of the recent initiatives which aims to fill this gap is the online Atlas of the Languages of Iran (ALI) (<http://iranatlas.net>). This article delineates objectives of the ALI research programme, atlas architecture, research methodology, and preliminary results that have been generated. Specific topics of interest are the structure and content of the linguistic data questionnaire; the handling of contrasting perspectives about the status of “languages” and “dialects” through a flexible multi-dimensional classification web; and the role of ongoing comparisons between language distribution assessments and hard linguistic data.

Erik Anonby is Visiting Researcher at the Leiden University Centre for Linguistics (the Netherlands), Associate Professor of French and Linguistics at Carleton University (Canada), and Visiting Professor in the Department of General Linguistics at the University of Bamberg (Germany).

Mortaza Taheri-Ardali is Assistant Professor of General Linguistics at Shahrekord University (Iran), and has been a Visiting Scholar at Radboud University (the Netherlands) and at the University of Bamberg (Germany).

Amos Hayes is Technical Manager at the Geomatics and Cartographic Research Centre at Carleton University (Canada).

The research activities and outputs featured in this article are the result of collaborative efforts by all members of the *Atlas of the Languages of Iran* (ALI) research team. Many colleagues have contributed to the development of ALI as well as to the writing of this article, but the requirements of this journal allow for only three authors. Special thanks go to Robert Oikle and Adam Stone for their contributions to the construction of the maps featured in this article. A full listing of institutional partners and individual ALI team members, along with their particular contributions to the research, is available at: <http://iranatlas.net/module/atlasteam>.

This work was supported by in-kind contributions from Atlas team members and their host institutions; a Carleton University Faculty of Arts and Social Sciences (FASS) Early Career Research Award; a Carleton University Development Award; a Social Sciences and Humanities Research Council (SSHRC) Insight Development Grant; and an Alexander von Humboldt Foundation Fellowship for Experienced Researchers.

Keywords: Iranian (Iranian) Languages; Languages of Iran; Turkic Languages; Language Documentation; Language Mapping; Language Classification; Atlas Design and Construction; Linguistic Data Questionnaire

Introduction

Throughout history, Iran has been situated at a crossroads of movements of peoples and intermingling of cultures. The rich linguistic heritage of the country reflects this diversity and dynamism. Many languages, from several language families, are represented: Iranian, Turkic, Semitic, Indic, Dravidian, Armenian, and Kartvelian, as well as sign languages. Some of the languages in these families have dozens, or even hundreds, of distinct dialects.¹ Yet the language situation of the country as a whole is still incompletely documented.

Since the 1950s, there have been a number of important efforts to map out the languages of Iran, but until now no language atlas, or even a comprehensive and detailed country-level language map, has been produced.

This paper opens with a review of important contributions to the mapping of Iran's languages, along with some recurring limitations and challenges. We then introduce one of the recent projects that seeks to provide a complementary, integrated picture of the language situation: the *Atlas of the Languages of Iran* (ALI) research programme (<http://iranatlas.net>).² The body of the article outlines the objectives of the ALI research programme, the architecture of the online atlas, research methodology, and preliminary results that have been generated. The article concludes with a discussion on the contribution of the atlas, and prospects for its ongoing development and enhancement.

History of Efforts

The concept of a language atlas for Iran is nothing new.³ Over sixty years ago, in the 1950s, the late Georges Redard at the University of Bern in Switzerland first moved ahead with the idea of working on an *Atlas Linguistique de l'Iran*. Data were gathered from 43 locations before the doors closed on the project, but these data have never been published.⁴

In the 1970s, the Persian Academy, in collaboration with the Iranian National Geographic Organization, likewise set out on an ambitious documentation project known

¹Both Persian and Kurdish, in the broad understanding of these labels, comprise hundreds of sub-groups with distinctive phonological, morphological, and syntactic profiles. For Persian in particular, Salāmi's (2004: 14) *Ganjineh-ye guyeshshenāsi-ye ostān-e Fārs* documents 72 (mostly Persian) dialects of one province in Iran (Fars Province), and this is only a small selection of the Persian dialects spoken there; in addition, several other provinces of Iran have large numbers of distinct Persian dialects, totaling several hundred across the country. For Kurdish, see Anonby, "Phonological Variation in Kurdish" and other chapters in the same volume.

²Anonby, Taheri-Ardali et al., *ALI = Atlas of the Languages of Iran*.

³This historical overview is a summary of Anonby, "Mapping Iran's Languages."

⁴Redard, "État des travaux," 7; Hāshemi, "Identification of Iran's Dialects," 166-7.

as *Farhangsāz*.⁵ This time, more than 16,000 questionnaires and 14,000 recordings were collected.⁶ However, with the change of political landscape at the end of the decade, the project was discontinued and it shared the fate of Redard's earlier work: with few exceptions,⁷ this wealth of data has never been analyzed or made available to the public.

The current successor to *Farhangsāz* is a large-scale language atlas programme which has been in operation for about seventeen years, and is carried out by the Iran Cultural Heritage Organization (ICHO).⁸ Some publications documenting the linguistic characteristics of particular regions and varieties have appeared,⁹ and further results in the form of an atlas, which is planned for publication in CD-ROM format,¹⁰ are eagerly awaited by the academic community and the public.¹¹

In addition to these country-wide atlas initiatives, a number of projects have been conducted which offer significant contributions to the mapping of Iran's languages, although the geographic and thematic scope of each of them is more specific or different from that of an atlas dedicated to the languages of Iran as a whole. Important examples of such atlas projects are: Stilo's *Atlas of the Araxes-Iran Linguistic Area*, which looks at linguistic features in specific locations of language contact and convergence in the north-western quarter of the country;¹² Labex EFL's planned mapping of linguistic microvariation in the Eastern Mediterranean–Iran–Caucasus region;¹³ and several linguistic map series¹⁴ and atlas projects initiated for specific regions of Iran.¹⁵ Key maps which show language distribution for the entire country, albeit as a generalized representation, are those of TAVO,¹⁶ Izady,¹⁷ Windfuhr,¹⁸ and Irancarto.¹⁹

In short, important language mapping research has taken place, and much more has been planned, but a comprehensive, detailed picture of the language situation in Iran,

⁵Modarresi, "Review: Iran's Dialects," 310.

⁶Ibid., 310; Mostapha Assi, Chair of Linguistics, IHCS (Institute for Humanities and Cultural Studies), pers. comm. 2015, in Anonby, "Mapping Iran's Languages."

⁷Subsequent to the termination of the *Farhangsāz* project, Mas'ud Purriāhi oversaw the production of several volumes on the languages of Esfahan Province using previously collected data. See, for example, Purriāhi and Foruzānfār, *Identification of Iran's Dialects*. In 2011, Sādeq Kiā's multi-dialect *Dictionary* was also published.

⁸Rostambeyk Tafreshi, "Hamadan Province," 61-2; Farāhāni Jam, "National Plan"; Parmun, *National Plan*.

⁹For example, Parmun and Zāre'i, *Language Atlas Fundamentals*.

¹⁰Faryār Akhlāqi, current project director, pers. comm. 2017.

¹¹Rostambeyk Tafreshi, "Hamadan Province," 61-2.

¹²Stilo, "Atlas of Araxes-Iran"; Stilo, "Araxes-Iran with Kurdophone Varieties."

¹³Labex EFL, "LC3."

¹⁴Borjān, *Continuity Between Persian and Caspian*; Stilo, "Isfahan Provincial Dialects."

¹⁵Among others, Asadpour, *Atlas-e zabāni-ye Āzarbāyjān-e Gharbi*; Deyhim, *Atlas-e guyeshshenākhtī-ye Qasrān-e Dākhel*; Doerfer, *Sprachgeographie des Chaladsch*; Rostambeyk Tafreshi, "Yazd Province"; Pāpoli-Yazdi, "Language Distribution Map." Irān Kalbāsi has supervised mapping work on many locations in Iran, for example, Bāziār, "Ilam Province."

¹⁶Orywal, *TAVO*, A.VIII.10.

¹⁷Izady, "Linguistic Composition of Iran."

¹⁸Windfuhr, "Dialectology and Topics," 11, 16.

¹⁹Hourcade et al., *Irancarto*.

and a systematic overview of key linguistic features of these languages, has yet to appear in the form of a fully-fledged atlas.

For those projects which aim toward this consequential goal, there have been, and remain, an array of impediments to the development and completion of an atlas of Iran's languages. These obstacles, which have been discussed more fully elsewhere,²⁰ include the sheer diversity of language families and individual language varieties spoken here; incomplete coverage and documentation of these languages, despite a history of documentation extending back more than a century;²¹ issues of funding, logistics and project design; contrasting perspectives on language identity and distribution;²² limited dissemination of project results; and limited cooperation among scholars working toward the common goal of a language atlas.²³ Any atlas project for Iran needs to come to terms with these issues in order to accomplish its objectives.

The ALI Research Programme

In order to address this gap in the literature, the *Atlas of the Languages of Iran* research programme was initiated in 2009, and first received seed funding in 2014. Today, ALI is an online, open-access resource (<http://iranatlas.net>), which is currently being developed by an international group of institutional partners and scholars.²⁴

Goals and guiding themes. The overall goal of the ALI research programme is to enable work toward a systematic understanding the language situation in Iran.²⁵ This initiative, which has the online Atlas at its core, is guided by a set of interrelated themes and questions:

- Linguistic typology: What are important linguistic features of Iran's languages and dialects, and how are they distributed geographically?
- Language distribution: Where are these language varieties spoken, and how does this compare to the distribution of linguistic features?
- Language classification: How do scholars and speakers classify these language varieties, and how can scholarly classifications be improved?
- Language documentation: A record of the linguistic situation in Iran and a repository of linguistic data before the disappearance of much—perhaps most—of the linguistic diversity with the extension of standard, Tehran-type Persian as a mother tongue across the country.

The ALI project team, research process and atlas architecture have been brought together and crafted in direct response to the challenges which face a project of this type, as outlined in the previous section ("History of Efforts").

²⁰Anonby, "Mapping Iran's Languages."

²¹Among others, Geiger and Kuhn, *Grundriss*; Zhukovskij, *Materialy*; Mann, *Mundarten*.

²²Anonby, Sabethematabadi, and Hayes, "Reconciling Contradictory Perspectives."

²³Anonby, "Mapping Iran's Languages."

²⁴A full listing of institutional partners and team members is available in Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/atlasteam>.

²⁵Anonby, "Overview of ALI Programme."

Research team. The ALI research group includes scholars from twelve countries, with Iran represented foremost among them.²⁶ The diverse constitution of the team ensures that a wide variety of opinions and expertise are represented and communicated to the Atlas' varied audiences, and volunteer contributions made by these researchers have enabled the project to move forward steadily.

Data platform, structure, and publication. ALI is published as an open-access, online resource and is being built by programmers at the Geomatics and Cartographic Research Centre (GCRC) using the open-source Nunaliit Atlas Framework (<http://nunaliit.org>),²⁷ which comes with a free and ready-made atlas template. In contrast to most commonly-used GIS packages such as ArcGIS or QGIS, which import static datasets into the program and focus on optimal visual representations, Nunaliit uses a "related-document" structure which is designed to handle complex relations among evolving data.²⁸ Inside a Nunaliit atlas, each piece of data is stored as a document with a flexible set of attributes, and each of these documents can be related to any other document in the atlas. This necessitates more initial set-up work in building an atlas, but once an atlas is operational, relations between data are easy to build, navigate, and process. For example, in a language atlas, an audio or video recording can easily be uploaded and associated with any other document (i.e. piece of data) in the atlas, such as:

- the particular settlement where it was collected, along with demographic information for that place;
- the category of linguistic data it shows: a certain word, grammatical construction, etc.;
- the language variety that it represents, and in turn, language classifications;
- clear acknowledgment of data sources: the speaker, and the person who contributed it to the atlas; or, in the case of published data, a reference to the website or book that it comes from.

All of these data types can be accessed directly through country- and province-level language distribution maps as well as other atlas modules including linguistic structure maps, language classification diagrams, and the bibliography.

The intricacy of the language situation in Iran necessitates a purposeful approach to data structure, and all of the country's some 60,000 cities and villages are placed in ALI as potential sites for data collection. Contrasting opinions on language classification add to this complexity. However, as will be shown below ("Language classification"), Nunaliit's related-document structure helps to define the geometry of relationships among language varieties and distinguish different kinds of relationships between them. Its interactive model-driven approach enables it to take into consideration

²⁶Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/atlasteam>.

²⁷GCRC, *Nunaliit Atlas Framework*; Hayes, Pulsifer, and Fiset, "Nunaliit Cybercartographic Atlas Framework."

²⁸Hayes and Taylor, "Developments in Nunaliit."

differing scholarly opinions about classification in the same diagram, and to produce different maps on the basis of these different perspectives.

Another key feature of a Nunaliit-designed atlas is its dynamic online platform. Nunaliit enables direct remote contributions by researchers, and by atlas users generally, from anywhere that has an internet connection, and collection and subsequent upload of data from locations without such a connection. To help ensure consistency and reliability, a system for moderation and double-checking of data is an integral part of the data contribution process. The simple fact that the atlas is online means that each portion—and even each piece of data—can be published as soon as it becomes available; there is no long wait for the release of a complete print version.

The open-access format of the Atlas, simultaneously published in Persian, English, and French, enables a wider, more engaged audience than a print version. As with other state-of-the-art atlases that use open-source code,²⁹ the programming work that goes into the production of each Nunaliit atlas is made freely available to other scholars on Github.³⁰

Considering all of these points, the basic choice of an online forum for ALI—and use of the Nunaliit Framework in particular—helps to address many of the challenges mentioned above (“History of Efforts”): an extremely complex language situation; contrasting perspectives on language classification and distribution; logistical considerations, including project funding; a forum for cooperation among scholars working toward the common goals of a language atlas; and dissemination of project results.

Key Atlas activities. The four main areas of current activity and production in ALI, proceeding from the programme objectives outlined above (“Goals and guiding themes”), are:

- linguistic structure maps;
- language distribution maps;
- language classification; and
- compilation of atlas resources.

This latter activity includes collection of oral texts as part of the fieldwork process, as well as the compilation of a reference list dedicated to materials relevant to mapping and documentation of the languages of Iran.³¹ This reference list is valuable because it points users to each piece of published data that is cited in the Atlas, and as it becomes more complete, it will serve as a comprehensive bibliography of all works related to the languages of Iran.

All of the four activities are essential to the development of a fully-fledged language atlas, and deeply integrated with one another through a cyclical research process in which each activity informs the others (for further discussion, see “Conclusion”):

²⁹For example, Dryer and Haspelmath, *WALS*; Michaelis et al., *APiCS*.

³⁰GCRC, “Developer Documentation.” <http://github.com/GCRC/nunaliit/wiki>.

³¹Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/bibliography>.

while linguistic structure maps may be viewed by linguists as central to a language atlas, and language distribution maps are often of greatest interest to a popular audience, language classification and compilation of atlas resources are seminal, ongoing activities that frame mapping activities and enable progress towards the production of maps. In the following sections, we will treat language classification, language distribution, and linguistic structure mapping in detail, and in this order, which reflects the methodological foundation and research process of the work.

Language Classification

In many language atlases, language classification is incidental to the research process. For Iran, however, careful treatment of this issue is essential to building a resource that handles linguistic data in a systematic way and is at the same time satisfactory for the diverse audiences of the Atlas.³² A comprehensive and well-grounded classification provides a context for identification of any language variety, review of scholarly consensus and differences of opinion, corroboration with perspectives of speakers and, ultimately, an overarching organizational structure for linguistic data in the Atlas. ALI's classificatory activities and features have been treated more fully elsewhere;³³ basic elements and highlights are reviewed here.

There are two models through which language classification is organized in ALI: a traditional classification organized as a two-dimensional tree structure,³⁴ and a multi-dimensional web model that allows for additional relational and representational possibilities.³⁵

In the two-dimensional classification, we bring together all language families, languages and dialects of Iran; and, in order to be topically complete, all languages of the Indic (Iranian) family which are spoken outside of Iran are also included. This working classification has been developed through consultation of the literature³⁶ and, as explained on the Atlas page dedicated to this two-dimensional model, the language families and subgroups of the tree have been reviewed by experts in the field.³⁷ A total of over 500 language varieties are inventoried here so far, of which about 400 belong to the Indic family. This count is constantly expanding as

³²Anonby and Sabethematabadi, "Complementary User Perspectives."

³³Anonby, "Three-Dimensional Approach;" Anonby et al., "Multi-Dimensional Approach."

³⁴Anonby, "Working Classification," in Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/classification>.

³⁵Ibid., <http://iranatlas.net/module/taxonomy.selectMap>.

³⁶Major sources for the classification of Indic are Stilo, "Tati Language Group," "Isfahan," "Caspian and Tati"; Lecoq, "Langues irano-aryennes occidentales," "Dialectes du sud-ouest," "Dialectes du centre"; Windfuhr, "Western Iranian Languages," "Dialectology," "Dialectology and Topics"; Skjærvø, "Documentation"; Borjian, "Median Dialects of Kashan," "Tabaroid"; Jahani and Korn, "Balochi"; and numerous contributions in Schmitt, *Compendium*. For Turkic, key sources are Bulut, "Turkic Varieties," "Convergence and Variation in Turkic." A full list of sources for these and other language families is found at the end of the traditional classification page (<http://iranatlas.net/module/classification>) and in the ALI bibliography (<http://iranatlas.net/module/bibliography>).

³⁷Anonby, "Working Classification," in Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/classification>.

ALI researchers encounter additional varieties in the literature as well as in the course of language distribution fieldwork.

One of the key features of this classification, as presented in the Atlas, is the backgrounding of assessments of a given variety as a language versus a dialect.³⁸ Many classifications,³⁹ as well as other types of linguistic and popular discourse, treat these two categories—"language" and "dialect"—as discrete and objectively identifiable. However, in the context of Iran, differences in the factors that people rely on for their assessments lead to greatly differing lists of languages for the country. Whereas official administrative materials usually specify between four and seven languages,⁴⁰ and popular sources have similar lists,⁴¹ scholars often cite dozens,⁴² and one source—*Ethnologue*—lists 80 distinct languages.⁴³ Counts given by speakers of the country's many languages are between the two extremes, but mother-tongue speakers of Tehran-type Persian consistently provide smaller lists than speakers of minority language varieties.⁴⁴ In short, there is no single, definitive list of Iran's languages that will adequately address the diverse audiences of the Atlas, even though all of them might expect one. But assessments of "language" vs. "dialect" do not contribute to the main purposes of the Atlas (see "Goals and guiding themes" above), and because of this we have chosen instead to focus on assembling an ever-expanding list of language varieties, and exploring the ways that these varieties fit together into a single coherent picture.⁴⁵

Because of its conceptual simplicity, informed by the comparative method,⁴⁶ the two-dimensional tree representation is useful as an indexing tool, and perhaps because of this, in practice it remains the dominant model in language classification today.⁴⁷ Still, even when distinctions of language and dialect are backgrounded, it is not ideally suited for the purposes of ALI, since contrasting viewpoints, arguments, and justifications for a given classification decision have to be relegated to footnotes and accompanying discussion.

³⁸Of course, the categories used in (scholarly or popular) language classifications may also include collective groupings such as phylum, language family, macrolanguage, and dialect group, as well as subgroupings such as subdialect, accent, sociolect, etc. However, the fundamental point of contention in the literature and in popular discourse is between varieties that are "real" languages and those that are "just" dialects.

³⁹For example, Simons and Fennig, *Ethnologue*; and Hammarström et al., *Glottolog*.

⁴⁰For example, SJS, *Census of Rural Districts*, as published in Hourcade et al., *Irancarto*, <http://www.irancarto.cnrs.fr/record.php?q=AR-040504&f=local&l=en> and <http://www.irancarto.cnrs.fr/record.php?q=AR-040503&f=local&l=en>;

⁴¹For example, Iran Traveling Center, "Languages in Iran", <http://www.irantravelingcenter.com/languages-in-iran>.

⁴²Orywal, *TAVO*, A.VIII.10.

⁴³Simons and Fennig, *Ethnologue*.

⁴⁴Anonby, Sabethemmatatabadi, and Hayes, "Reconciling Contradictory Perspectives"; Sabethemmatatabadi, Anonby, and Sheyholislami, "Language Classification Taxonomies."

⁴⁵*Ibid.*; Anonby, "Three-Dimensional Approach"; Anonby et al., "Multi-Dimensional Approach."

⁴⁶Leskien, *Declination*.

⁴⁷Anonby, "Three-Dimensional Approach"; Anonby et al., "Multi-Dimensional Approach."

Despite its pervasiveness in historical linguistics, the tree model of classification has come under persistent criticism⁴⁸ since shortly after its elaboration by Schleicher in the mid-1800s.⁴⁹ Among other issues, it overlooks the contribution of areal phenomena such as transitional structures and effects of contact, whatever the genealogical (i.e. genetic) relation⁵⁰ between language varieties. A “wave” model of language relationship, proposed in response to this gap,⁵¹ has increasingly gained traction among scholars; and sophisticated expressions of the wave model, informed by the comparative method, have been developed recently.⁵²

For issues of classification, the ALI research team seeks to handle fundamental constraints of existing models, treat the complexity of the language situation in a reasoned and systematic way, and to operationalize new insights in the design of the Atlas. To this end, the authors of this article, assisted by geographic data technicians and programmers at the GCRC, have developed a multi-dimensional web model of classification for the languages of Iran (Figures 1 and 2).⁵³ This model contains all of the same content as the two-dimensional tree model, but allows for various types of relations to be expressed within the same diagram. The labelled diagram in Figure 2, while visually overwhelming, is a fitting representation of the complex language situation.

Genealogical relations, as established by the comparative method, form a central organizing principle for the representation (Figure 3). Structural similarity through language contact is another way in which varieties can be connected to one another (Figure 4), as is relationship through ethnic identification (Figure 5). Explanations and examples of each of these types of relations are more fully detailed in the studies dedicated to this topic.⁵⁴

While the multi-dimensional model for language classification has wider implications for historical linguistics, it also provides a reasoned and systematic way of making explicit how the complexities of the language situation are treated in the Atlas. Importantly, the ALI project is currently exploring the way in which presuppositions and perspectives from different audiences, when applied systematically, define the maps that are produced.⁵⁵ Ongoing refinement of language classifications through

⁴⁸See especially Schmidt, *Verwandschaftsverhältnisse*; Aikhenvald and Dixon, *Areal Diffusion*; Hegarty, “Prehistory Through Language”; François, “Models of Language Diversification.”

⁴⁹Schleicher, “Indogermanischen Urvolkes.”

⁵⁰Following Haspelmath, “Genealogical Linguistics,” 222, we use the term “genealogical” in place of the traditional term “genetic” to avoid confusion with biological genetic relationships.

⁵¹Schmidt, *Verwandschaftsverhältnisse*.

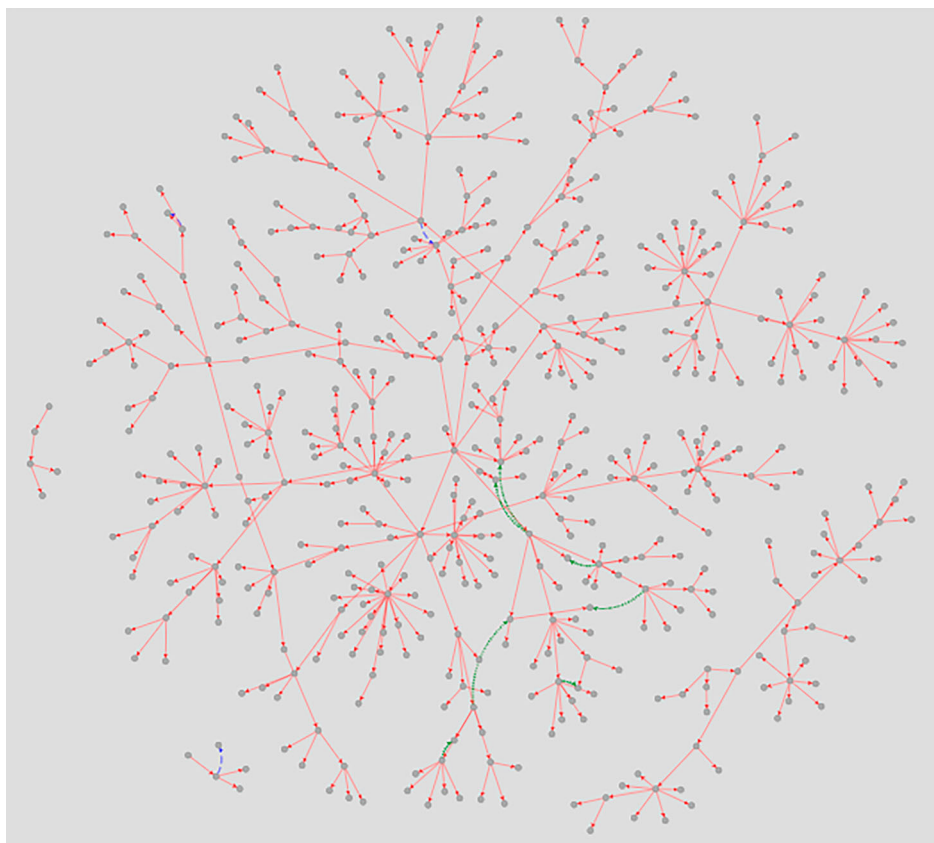
⁵²François, “Models of Language Diversification”; Kalyan and François, “Freeing the Historical Method”; Jügel, “Linguistic History of Kurdish.”

⁵³Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/taxonomy.selectMap>.

⁵⁴Anonby, “Three-Dimensional Approach”; Anonby et al., “Multi-Dimensional Approach.”

⁵⁵To give a well-known example in relation to Kurdish: people who treat ethnic identification as a starting point might show Kurdish on a map as a single language, but people who base their work on genealogical relations, as established by the comparative method, might show three or more distinct languages: Kurdish (itself possibly further subdivided into Northern, Central and Southern Kurdish), Zazaki, and Gorani. This and other examples are discussed further in Anonby et al., “Multi-Dimensional Approach.”

Figure 1. The Languages of Iran as Represented in a Multi-Dimensional Language Classification Web.



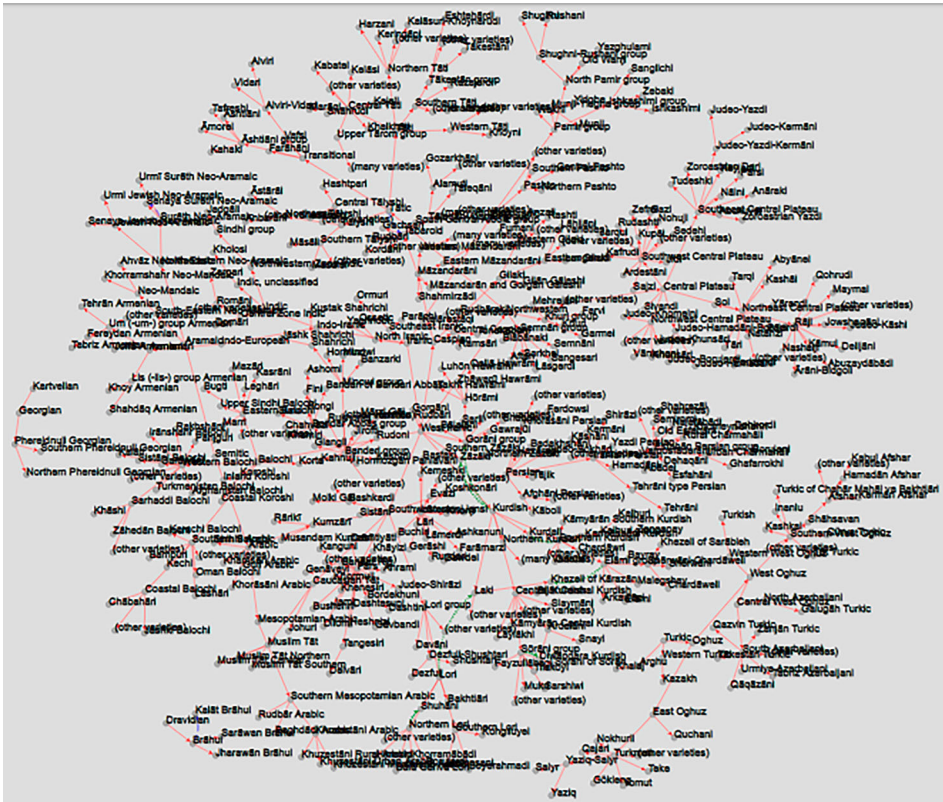
Source: <http://iranatlas.net/module/taxonomy.selectMap>

field research and analysis of linguistic data, discussed in the following sections, sets in motion a dialogical research process for the Atlas.

Language Distribution and Local Place Names

While linguistic structure maps—described in the following section—are an ultimate goal for the output of the Atlas, there is a significant amount of preparatory work that goes on for each province before linguistic data collection can take place. Since it is not practical (and probably not even possible) to gather systematic linguistic data from each language community in each of Iran's some 60,000 cities and villages, even with an extended time frame, the research team needs to select research sites judiciously and economically. This, in turn, requires familiarity with

Figure 2. The Languages of Iran in a Multi-Dimensional Language Classification Web, with Language Variety Labels. *Note:* this figure shows static output from the interactive module.



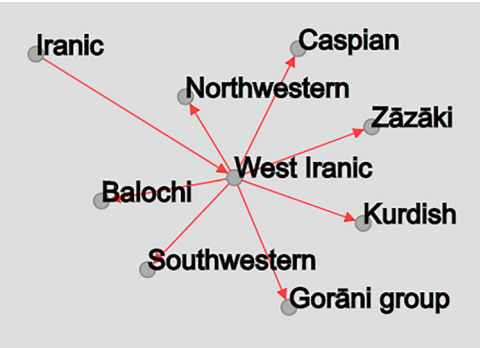
Source: <http://iranatlas.net/module/taxonomy.selectMap>

the nature and distribution of all of the language varieties that are spoken in each province.

In a series of recent papers, we provide detailed case studies of the language distribution research process for three provinces: Kordestan, Chahar Mahal va Bakhtiari, and Ilam.⁵⁶ We are committed to establishing a streamlined, replicable research process for the study of language distribution, but the exact implementation varies from one province to another. It depends on the availability, back-

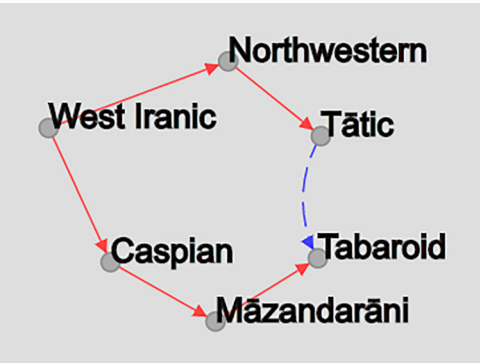
⁵⁶Aliakbari, Gheitasi, and Anonby, “Language Distribution in Ilam”; Anonby, Gheitasi, and Aliakbari, “Revisiting Language Distribution”; Anonby, Mohammadirad, and Sheyholislami, “Kordestan Province”; Taheri-Ardali and Anonby, “Language Distribution”; Anonby, Taheri-Ardali, and Stone, “Linguistic Variation in Chahar Mahal va Bakhtiari.”

Figure 3. Sample Multi-Dimensional Representation of Genealogical Relations.



Source: <http://iranatlas.net/module/taxonomy.selectMap>

Figure 4. Sample Multi-Dimensional Representation of Structural Similarity through Language Contact.

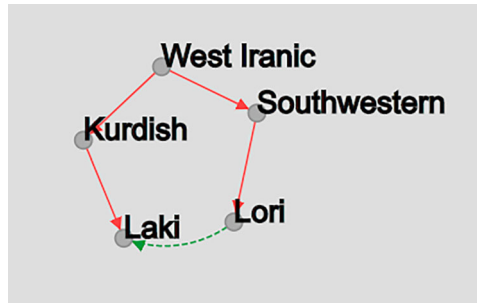


Source: <http://iranatlas.net/module/taxonomy.selectMap>

ground, and expertise of the scholars who will carry out this research and, of course, fieldwork conditions on the ground. While language distribution profiles and maps for some provinces have been completed after several months of intensive work, other provinces have taken more than two years to realize.

As mentioned in the paragraphs on “Data platform, structure and publication” above, substantial work goes into the preparation and structuring of the language distribution maps, because of the complexity of the data. A first step for each province is the compilation of a bibliography of all resources pertaining to the distribution of languages of that province, and all resources that contain linguistic data from those languages. From these materials, we are able to establish an initial list of all language varieties that we expect to encounter. With the help of linguists who have worked on

Figure 5. Sample Multi-Dimensional Representation of Relationship through Ethnic Identification.



Source: <http://iranatlas.net/module/taxonomy.selectMap>

the languages of the region, we organize the varieties into an initial working classification—an initial hypothesis that is scrutinized and tested in the subsequent fieldwork.

We then bring together geographic data⁵⁷ and demographic data⁵⁸ which the government of Iran makes publicly available. These data are imported and associated with a background map for each province, specially designed to orient the user and foreground language-related data (Figure 6). The background map, with all associated geographic and demographic data, is then ready to be populated by language distribution data gathered through fieldwork.

Each province is managed by at least one lead researcher, primarily responsible for that area, who directs language distribution fieldwork and carries it out either in conjunction with a team of colleagues (whether senior scholars or students), or independently. For every city or village in each province, the researchers investigate two general topics: language distribution, and local place names.

First, as the major issue, language distribution: (a) What languages are spoken in this settlement as a mother tongue?; (b) Which sub-varieties of these languages are found here?; and (c) What is the estimated proportion of people who speak each variety as a mother tongue?

Members of the research team start out with a solid, fundamental understanding of the language situation in the province they are studying. However, because of the settlement-level scale of the research, in each case they have made important discoveries for themselves regarding the diversity and regional distribution of language varieties. Simply by asking these questions in relation to each settlement, researchers frequently encounter language varieties that have not been reported in the literature; and the local context of their research brings to light language communities' own perspectives on how their language fits into a more general

⁵⁷National Cartographic Center, *Database of Geographic Names*; Rural Network of Iran, *Roostanet*.

⁵⁸Iran Statistics Center, *Public Census*.

Figure 6. Sample Background Map: Hormozgan Province.



Source: <http://iranatlas.net/module/language-distribution.hormozgan>

taxonomy of languages in the country. Both of these elements feed into the elaboration and refinement of ALI’s comprehensive working classification (see “Language Classification” above).

Results of the language distribution research are indicated in each relevant place document, along with the name of the researcher who has contributed their assessment (Figure 7). This stage of the research process, described in detail elsewhere,⁵⁹ is clearly not a census of all individuals, nor would it be feasible to aim for this. Rather, these investigations yield estimates for each place and are clearly indicated as such. Linguists and other Atlas users who have more detailed local knowledge about language distribution in particular settlements are able to propose corrections and updates directly through the moderated Atlas interface (see “Data platform, structure and publication” above).

At the same time as investigating language distribution, researchers ask about local names for each settlement, as pronounced in the languages spoken there. Often, the local name is different—whether partially or completely—from the official Persian name. A few examples from Bushehr Province are as follows:⁶⁰

<u>Official name</u>	<u>Local name</u>
Berikān	Biriku

⁵⁹Anonby, Mohammadirad, and Sheyholislami, “Kordestan Province.”

⁶⁰Fatemeh Nemati and Shakiba Ghasemi, “Bushehr Province,” in Anonby, Taheri-Ardali et al., *ALI*, <http://iranatlas.net/module/language-distribution.bushehr>.

Figure 7. Sample Language Distribution Document: Qorveh.

Qorveh <i>city</i>	
Population	71232
Location (lat, long)	35.16735, 47.80384
Province (roman)	Kordestān
Shahrestān (roman)	Qorveh
Bakhsh (roman)	Markazi
City (roman)	Qorveh
Local name	Qurwa
Local name source	Masoud Mohammadirad, field notes 2015
Language distribution estimate	Tehrāni type Persian 10% Turkic 15% Central Kurdish 54% Southern Kurdish 18% Hōrāmi 3%
Language distribution source	Masoud Mohammadirad, field notes 2015
Atlas unique ID	121011

Source: <http://iranatlas.net/module/language-distribution.kordestan>

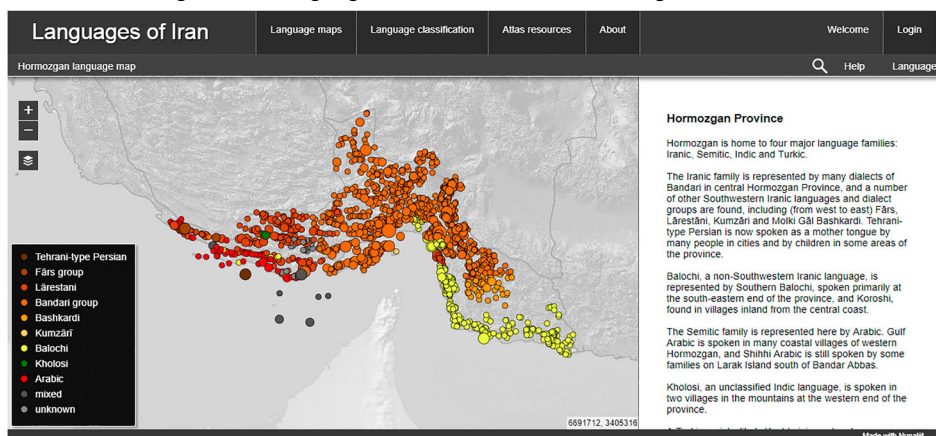
Fāryāb	Pariyow
Mohammad Shāhi	Momshey
Tombak	Aynak
Ahshām-e Hasan	Kheshem-e Kalsan

Local pronunciations are an important means of highlighting the diversity of the language situation on the ground, and serve as a way of connecting with each community featured in the Atlas. Phonemic transcriptions⁶¹ are supplied for each place, and where the research team has prepared them, IPA (International Phonetic Alphabet) transcriptions and/or recordings are also made available.

Now three years into the ground-laying process of language distribution research, interactive point-based language distribution maps in ALI have been completed for five provinces of Iran, in the following order: Hormozgan, Kordestan, Chahar Mahal va Bakhtiari, Bushehr, and Ilam (Figures 8–12).

⁶¹“Transcription Conventions,” in Anonby, Taheri-Ardali et al., *ALI*, <https://carleton.ca/iran/transcription>.

Figure 8. Language Distribution in Hormozgan Province.



Source: <http://iranatlas.net/module/language-distribution.hormozgan>

We are also developing the capability of constructing country-level language distribution maps for any individual variety or group of varieties selected by Atlas users (Figure 13).⁶²

Finally, in response to requests from other scholars, we have generated traditional static polygon maps as an alternative representation of language distribution (Figure 14).⁶³ These three types of maps constitute the first series of fine-grained language distribution maps published for any part of Iran, and together they are contributing to a more coherent and detailed picture of language distribution for the country as a whole.

Importantly, for the purposes of the ALI research programme, the results of language distribution research also provide a basis for representative and economical selection of research sites for collection of linguistic data. This next stage of the research is described in the following section.

Collection and Mapping of Linguistic Data

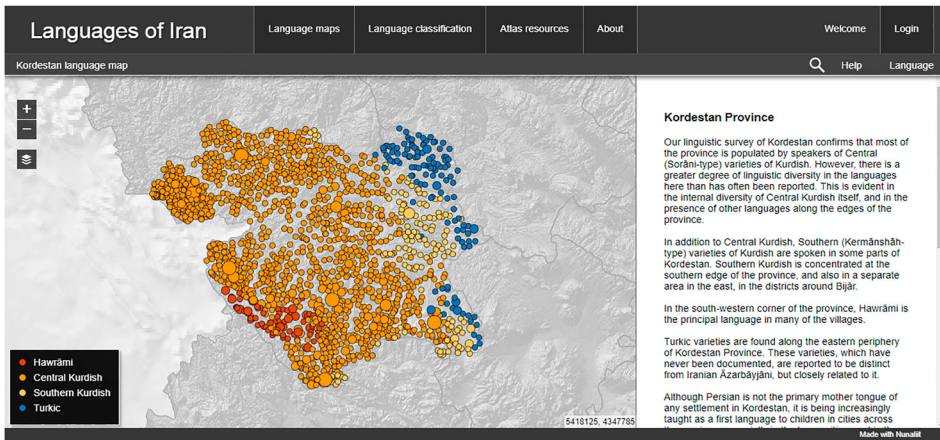
The mapping of linguistic data—words, phonological features, grammatical features—is at the heart of any fully-fledged language atlas. All of the elements described so far, and especially research on language classification and language distribution, are necessary precursors to the collection, analysis, and mapping of linguistic data, since these data are principally meaningful in relation to their wider comparative linguistic and social context.

The language distribution maps developed for each province provide an accessible point of reference for identifying evenly distributed sites for collection of linguistic

⁶²“Distribution of Individual Languages and Language Families,” in Anonby, Taheri-Ardali et al., *ALI*, http://iranatlas.net/module/language-distribution.single_language; Anonby and Sabethemmatbadi, “Complementary User Perspectives.”

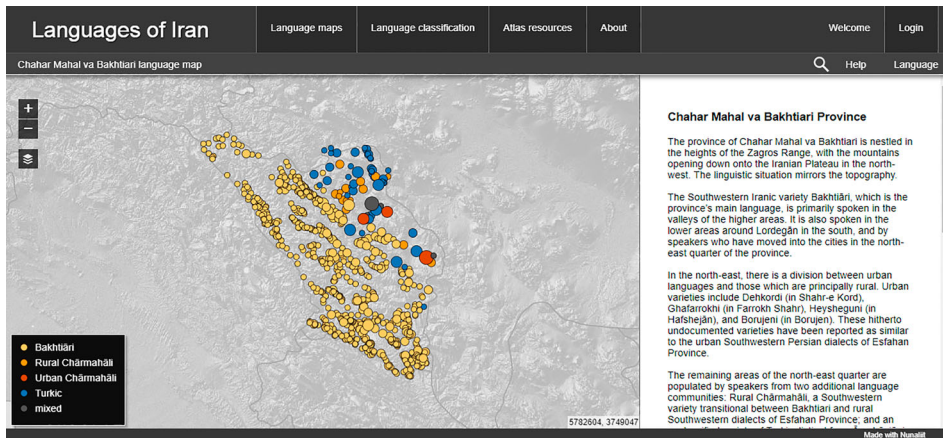
⁶³Anonby, Taheri-Ardali et al., *ALI*. These maps are available from the homepage by selecting “Language maps” > “Sample linguistic data maps” > “Chahar Mahal va Bakhtiari” from the menu.

Figure 9. Language Distribution in Kordestan Province.



Source: <http://iranatlas.net/module/language-distribution.kordestan>

Figure 10. Language Distribution in Chahar Mahal va Bakhtiari Province.

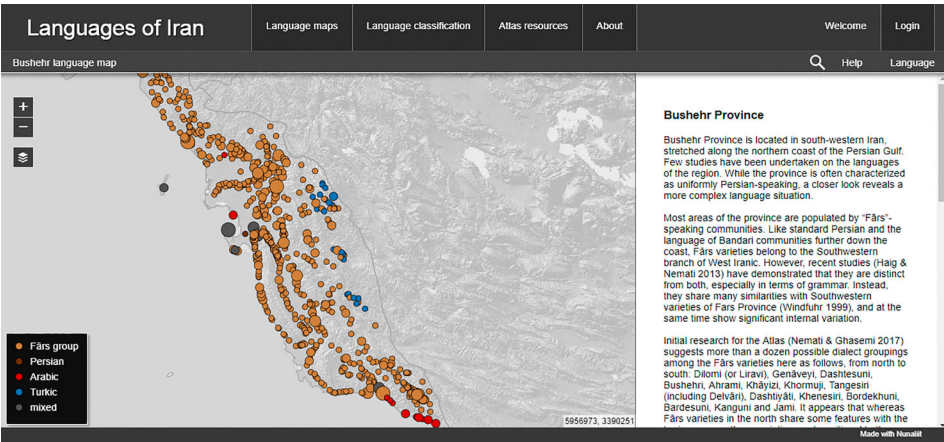


Source: http://iranatlas.net/module/language-distribution.chahar_mahal_va_bakhtiari

data. Generally, for each language variety⁶⁴ identified in the language distribution phase, the Atlas aims to collect linguistic data from one representative site in each *shahrestān* (provincial sub-district) where that variety is found. In addition, since

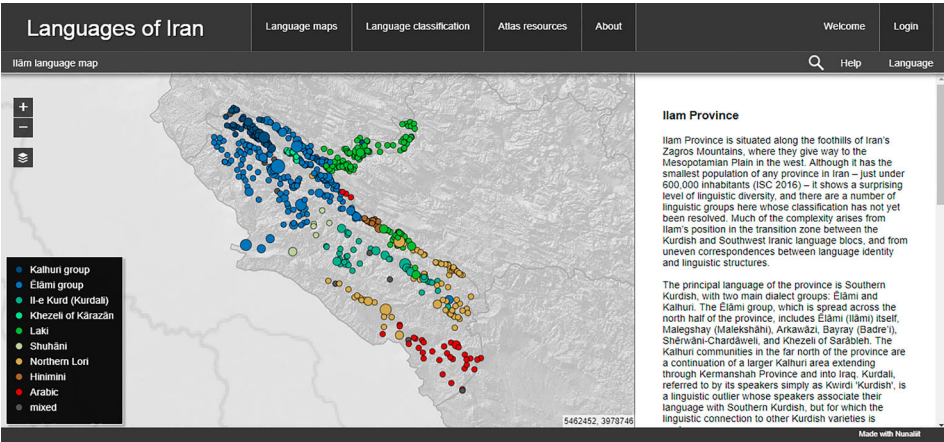
⁶⁴That is, each language variety grouping which is identified by speakers above the level of single-location varieties (unless a single-location variety is indicated by speakers and other people in a region as distinctive, in which case it is also included).

Figure 11. Language Distribution in Bushehr Province.



Source: <http://iranatlas.net/module/language-distribution.bushehr>

Figure 12. Language Distribution in Ilam Province.

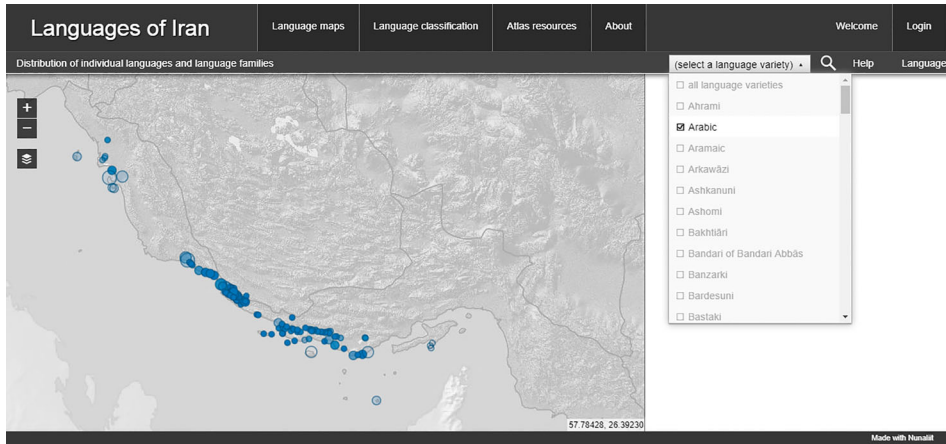


Source: <http://iranatlas.net/module/language-distribution.ilam>

language contact phenomena are an important aspect of the language situation, data collection sites are selected at the edge of language areas, in dialectically transitional areas, and in multi-lingual cities and villages.

For larger, more linguistically diverse provinces like Hormozgan (Figure 8), these parameters necessitate data collection from well over 50 research sites. Even for a small and relatively less diverse province such as Chahar Mahal va Bakhtiari, we are aiming to collect data from thirty-five locations. At present, the research team for that province, led by Mortaza Taheri-Ardali, has collected ALI pilot questionnaire

Figure 13: Sample Language Distribution Map for Individual Varieties: Arabic in Bushehr and Hormozgan.



Source: http://iranatlas.net/module/language-distribution.single_language

data from thirty of these locations (Figure 15).⁶⁵ As Figure 15 shows, the selected linguistic data collection sites are clustered in districts with a higher level of linguistic diversity.

Other sites where the Atlas team has already carried the pilot questionnaire out are located in Hormozgan (10 locations), Ilam (5), Kermanshah (1), Esfahan (1), Markazi (1), Boyerahmad va Kohgiluyeh (1) and Sistan va Baluchistan (1), bringing the total to 50 locations so far.

Two types of data are collected in each location: oral texts, including high-quality audio as well as video format whenever possible; and a uniform dataset across all research sites based on responses to the ALI linguistic data questionnaire.

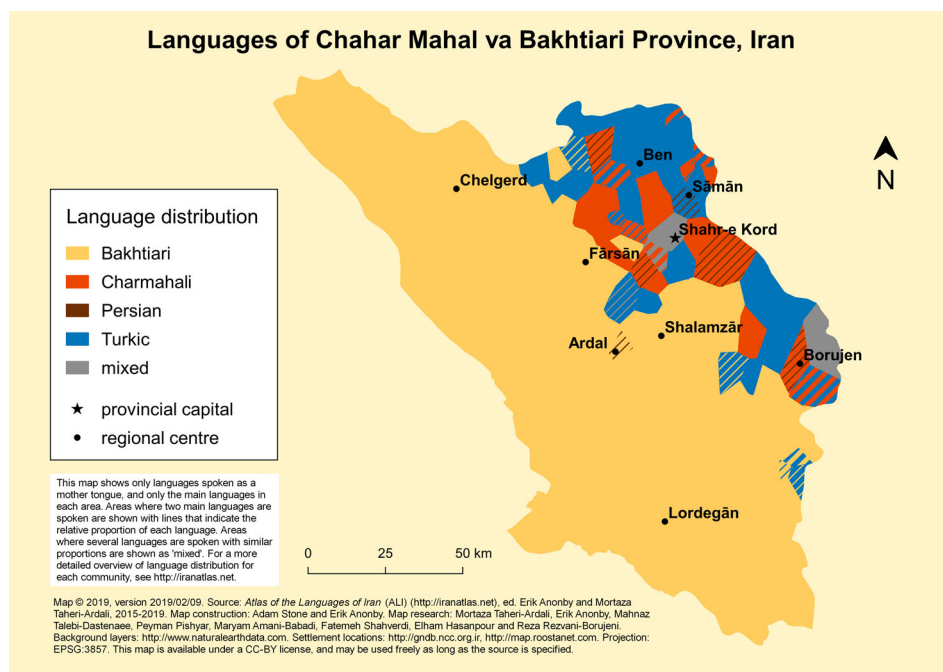
The ALI linguistic data questionnaire. Collection and analysis of oral texts is essential in understanding natural patterns of discourse and language use; in contrast, design and administration of a linguistic data questionnaire is a systematic and efficient means of gathering focused data that can answer very basic—but hitherto unresolved—questions about the language situation in Iran, as well as the typological diversity and geographic distribution of linguistic structures.

The ALI questionnaire is designed specifically for the languages of Iran, with attention to “classic” features and isoglosses identified and applied by scholars in the field of Iranian linguistics, but going beyond these as well. Along with recurrent themes in the literature such as those found in the *Compendium Linguarum Iranicarum*,⁶⁶ the fea-

⁶⁵Taheri-Ardali and Anonby, “Language Distribution”; Anonby, Taheri-Ardali, and Stone, “Linguistic Variation in Chahar Mahal va Bakhtiari.”

⁶⁶Schmitt, *Compendium*.

Figure 14. Static Polygon Representation of Language Distribution in Chahar Mahal va Bakhtiari Province.



Source: http://iranatlas.net/module/language-distribution.chahar_mahal_va_bakhtiari_static

tures treated in Stilo's *Araxes-Iran* atlas project⁶⁷ and the dialectological questionnaire published by the Persian Academy (*Farhangestān*)⁶⁸ have been foundational in defining the content of the ALI questionnaire. In order to ensure typological breadth and applicability to other language families in Iran, the questionnaire has been further developed in consultation with, among others, Comrie and Smith's (1977) typological questionnaire, still a standard in the field today;⁶⁹ features in the *World Atlas of Language Structures* (WALS);⁷⁰ and the Swadesh and Leipzig-Jakarta wordlists.⁷¹ Chan's questionnaire on numeral systems⁷² has also been incorporated.

The main elements of the questionnaire, for which parallel Persian and English versions are available, are: sociolinguistic context for each research site; lexicon; phonol-

⁶⁷Stilo, "Atlas of Araxes-Iran"; Stilo, "Araxes-Iran with Kurdophone Varieties."

⁶⁸Persian Academy, "Rāhnemā."

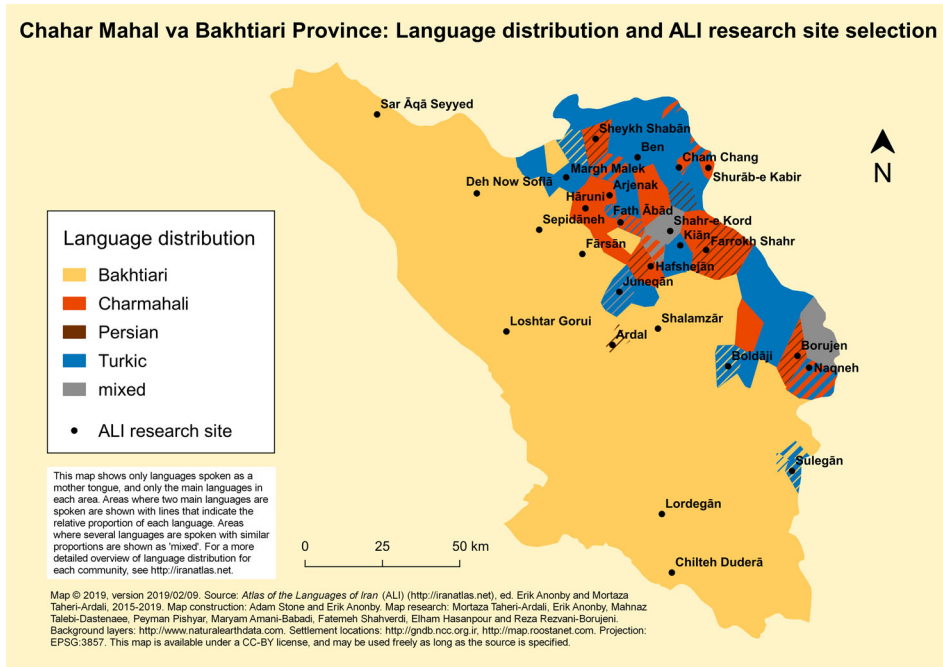
⁶⁹Comrie and Smith, "Lingua Descriptive Studies Questionnaire."

⁷⁰Dryer and Haspelmath, *WALS*.

⁷¹Swadesh, *Origin and Diversification*, 283; "Leipzig-Jakarta List of Basic Vocabulary", in Haspelmath and Tadmor, *Loanwords*, 69-71.

⁷²Chan, *Numeral Systems*.

Figure 15. Locations for Collection of Linguistic Data in Chahar Mahal va Bakhtiari Province.



Source: <http://iranatlas.net/module/linguistic-data.cb-research-sites>

ogy; and morphosyntax (grammar).⁷³ Over a two-year period, from mid-2015 to mid-2017, the pilot version of this questionnaire was carried out in the fifty locations mentioned earlier in this section.

In the summer of 2017, the questionnaire was revised with the in-depth input of experts as part of a workshop at Bamberg University.⁷⁴ Major improvements to the questionnaire, according to the various topics covered, were as follows:

- Sociolinguistic context: inclusion of questions that help provide a more detailed picture of respondents' linguistic background and use, since these factors have a major impact on the linguistic structures that are elicited.
- Lexicon: augmentation of the lexicon section, which was straightforward to administer, and which produced clear and meaningful results, from 80 to 130 words; additional elicitation supports, such as explanations and pictures, for semantically problematic items.

⁷³Anonby, Taheri-Ardali et al., *ALI*, <https://carleton.ca/iran/questionnaires>.

⁷⁴Workshop on Refining the questionnaire for the Atlas of the Languages of Iran, July 2017, <https://www.uni-bamberg.de/aspra/workshop-questionnaire-languages-of-iran-2017>.

- Phonology: removal of the dedicated phonology section, including the comprehensive phonemic inventory, which required a long segment of interview time, a high level of phonological expertise, and the ability to design and conduct ad hoc analyses and proofs; redistribution of the more basic and important synchronic as well as historical phonology-related questions into the lexicon and morphosyntax sections.
- Morphosyntax: complete redesign of the morphosyntax section to enable researchers without expertise in this domain, or in a given language family, to administer the questions more easily; provision of supplementary context for discourse-dependent topics such as definiteness and morphosyntactic alignment; elicitation and tagging of data according to grammatical function⁷⁵ rather than a closed set of expected structural possibilities.

In addition, the questionnaire as a whole was repurposed to take into account considerations of design and content related to Turkic, Semitic, and the other language families found in Iran.

Following on testing of this refined, definitive version of the questionnaire,⁷⁶ it has been recently released for use in the ALI project.⁷⁷

Processing and mapping of linguistic data. Whenever field researchers complete linguistic data collection using the ALI questionnaire, they transcribe all the data according to the project transcription conventions⁷⁸ and send it to the province team manager for an initial check. Once outstanding questions have been defined and addressed, the manager then brings together and organizes, in an Excel file, data from all research sites in that particular province. The Atlas editors carry out a further check of the data and, together with all colleagues who have worked on the data, identify regularities in the linguistic patterning and geographic distribution of individual structures.

Results of linguistic data collection are currently being analyzed and released as papers⁷⁹ and as maps on the Atlas website. The region for which data collection, processing, and publication is most advanced is that of Chahar Mahal va Bakhtiari Province, where comparative analysis has been carried out (Figure 16 shows a sample of the data) and a series of static lexical and phonological maps have been produced (Figures 17–19). While these maps reveal meaningful patterns on their own, they are important as a foretaste of the contributions that will be

⁷⁵Haig and Schnell, “Annotations Using GRAID.”

⁷⁶Typology and Documentation Workshop, Hamadan University, February 2018.

⁷⁷Anonby, Taheri-Ardali et al., *ALI*, <https://carleton.ca/iran/questionnaires>.

⁷⁸“Transcription Conventions,” in Anonby, Taheri-Ardali et al., *ALI*, <https://carleton.ca/iran/transcription>.

⁷⁹Anonby and Schreiber, “Qashqay Turkic of Shiraz”; Anonby and Taheri-Ardali, “Bakhtiari”; Anonby and Taheri-Ardali, “Chahar Mahal va Bakhtiari”; Schreiber et al., “Contact-induced Change in Turkic”; Schreiber et al., “Variation in Qashqay Turkic”; and Taheri-Ardali, “Languages in Western Hormozgan.”

Figure 16. Sample Lexical Data Sets from Chahar Mahal va Bakhtiari Province.

		'wood'		'big'	
B	Sar Aqā Seyyed	ču		gahp	
B	Deh Now Soflā	ču		gahp	
B	Sepidāneh	ču		gahp	
B	Fārsān	ču		gahp	
B	Loshtar Gorui	ču		gahp	dalū
B	Ardal	ču		gahp	
B	Shalamzār	ču		gahp	gonde
B	Boldāji	čub	ču	gonde	
B	Lordegān	ču		gapu	
B	Chilteh Doderā	ču		gapu	
C	Fath Abād	čub	čug	gonde	bozorg
C	Cham Chang	čug		gonde	bozorg
C	Shurāb Kabir	ču		gahp	gonde
C	Sheykh Shabān	čug	čub	gonde	
C	Arjenak	čug		gonde	
C	Haruni	čug		bozorg	gonde
C	Shahr-e Kord	čug	čub	gonde	bozorg
C	Farrokh Shahr	čug		gonde	
C	Hafshejān	čug		gap	bozorg
C	Borujen	čug		gonde	
C	Naqneh	čug		gonde	bozorg

made by maps in other domains—notably morphosyntax—and in other provinces of Iran. Currently, as the data are being analyzed, it is giving insight into planning for the kind of programmatic structuring that will be necessary to turn the Atlas' experimental dynamic language maps into fully developed maps—maps in which the form- and function-tagged linguistic data are fully integrated with the existing architecture of the Atlas and made accessible to users for easy output of the data in tabular format and creation of their own interactive maps in the Atlas.

Conclusion

This article provides a first global description of the *Atlas of the Languages of Iran* (ALI) research programme. It begins with an account of efforts to map the languages of Iran, from the 1950s until today, and a review of challenges that have faced this enterprise.

The ALI programme, designed to address these challenges, was initiated in 2009 and has moved forward steadily since seed funding was obtained in 2014. The overarching purpose of this research programme is to enable work toward a systematic

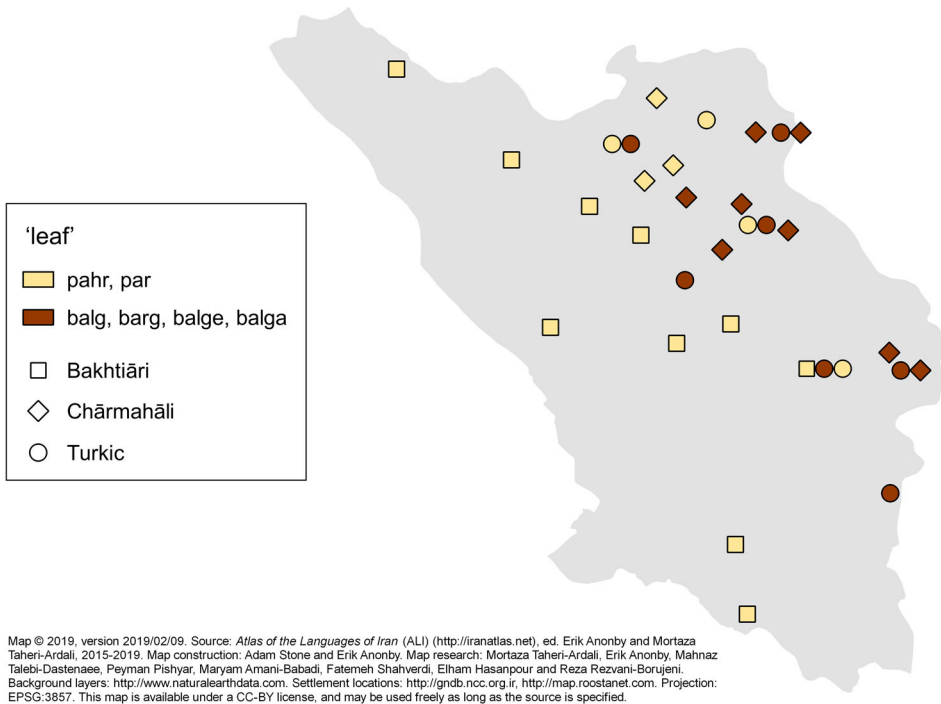
understanding the language situation in Iran. This work, which has an online Atlas at its core, revolves around the exploration of the typology, distribution, and classification of the languages of Iran. Further, Atlas activities are conceived to contribute to the documentation of the language situation in the face of the decline and, in many cases, disappearance of these languages as standard, Tehran-type Persian is increasingly adopted as a mother tongue in areas across the country.

The Atlas at the centre of this project is built using the Nunaliit Atlas Framework. Nunaliit's related-document approach and a range of rendering and input options enable a research process and product capable of handling complexities of the language situation in Iran and reaching the diverse audiences of the Atlas in an accessible way. Key elements of present Atlas production include a multi-dimensional working classification of all language varieties in the country; language distribution maps; and linguistic structure maps generated through the collection and analysis of systematic questionnaire data. While all of these elements are important in their own right, each one contributes to the understanding and development of the others. Ongoing comparison of linguistic structure maps with language classifications, and with language distribution maps, ensures that the picture that emerges is grounded in the actual forms that speakers use. This multi-faceted perspective is critical, since linguistic data are often ambivalent with respect to language classification, and rarely show neat correspondences to perceived geographic boundaries between language varieties. Conversely, a clearly elaborated language classification and a detailed picture of language distribution provide a comparative linguistic, and social, context for these data.

Areas of current activity include the study of language distribution for remaining provinces of Iran; carrying out of the revised linguistic data questionnaire in provinces where language distribution research has been completed; and construction of an accessible portal where linguistic data are made available for download and display on user-constructed maps. Thanks to the open-source technology and the dedication of the research team, the project is sustainable in a context of modest funding. Building on the foundation of more than a century of documentation, and alongside complementary initiatives, the *Atlas of the Languages of Iran* is contributing to a clearer picture of the diverse language situation in Iran.

Figure 17. Sample Lexical Data Map: ‘Leaf’ in the Languages of Chahar Mahal va Bakhtiari Province.

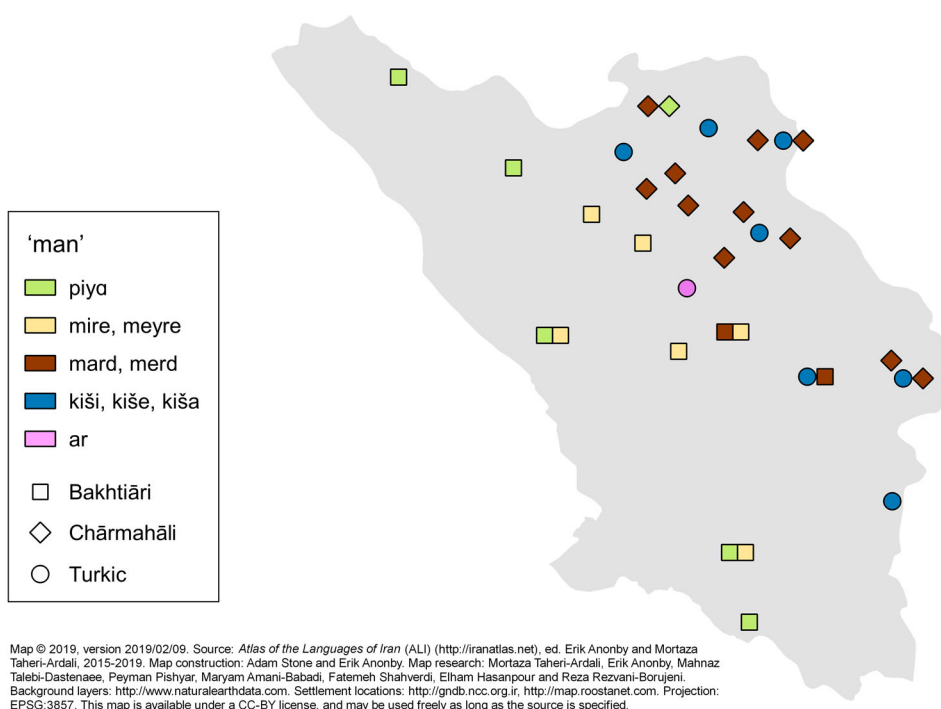
Lexical variation in Chahar Mahal va Bakhtiari Province: ‘leaf’



Source: <http://iranatlas.net/module/taxonomy.selectMap>

Figure 18. Sample Lexical Data Map: “Man” in the Languages of Chahar Mahal va Bakhtiari Province.

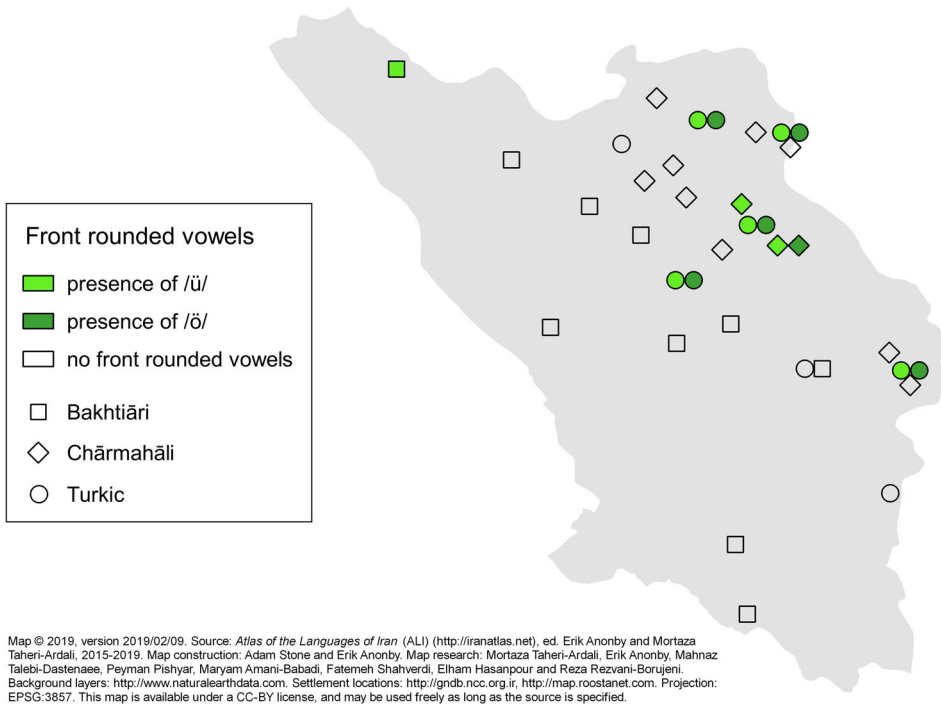
Lexical variation in Chahar Mahal va Bakhtiari Province: ‘man’



Source: <http://iranatlas.net/module/taxonomy.selectMap>

Figure 19. Sample Phonological Data Map: Front Rounded Vowels in the Languages of Chahar Mahal va Bakhtiari Province.

Phonological variation in C&B Province: Front rounded vowels



Source: <http://iranatlas.net/module/taxonomy.selectMap>

Bibliography

- Afrāshi, Āzītā. "Guyesh-e Komzāri-ye jazireh-ye Lārak: Āzmun-e melākhā-ye radehshenākhti" [The Kumzari dialect of Larak Island: testing typological criteria]. *Majjaleh-ye zabān-ha va guyeshbā-ye Irāni (Farhangestān-e zabān va adab-e fārsi)* [Journal of Iranian Languages and Dialects (Academy of Persian Language and Literature)] (new series) 1 (1391/2012): 139-60.
- Aikhenvald, Alexandra Y., and R.M.W. Dixon, eds. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press, 2001.
- Aliakbari, Mohammad Aliakbari, Mojtaba Gheitasi and Erik Anonby. "On Language Distribution in Ilam Province, Iran." *Iranian Studies* 48, no. 6 (2014): 1-16.
- Anonby, Erik. "Phonological Variation in Kurdish." In *Variation and Typology in the Dialects of Kurdish*, edited by Yaron Matras, Geoffrey Haig and Ergin Öpengin. London: Palgrave Macmillan, forthcoming.
- Anonby, Erik. "Mapping Iran's Languages: Situation and Prospects." Paper presented at the 6th International Conference of Iranian Linguistics (ICIL6), Ilia State University, Tbilisi, Georgia, June 23-26, 2015.
- Anonby, Erik. "Overview of the ALI (*Atlas of the Languages of Iran*) Research Programme." Paper presented at the Workshop on Refining the ALI Questionnaire, Universität Bamberg, July 7-8, 2017.
- Anonby, Erik. "A Three-Dimensional Approach to Classification of Iran's Languages." Paper presented at the 1st North American Conference in Iranian Linguistics (NACIL1), Stony Brook University, New York, April 28-30, 2017.
- Anonby, Erik, Mojtaba Gheitasi, and Mohammad Aliakbari. 2017. "Revisiting Language Distribution in Ilam Province." Paper presented at the 3rd International Conference on Kurdish Studies (ICKS3), University of Exeter, June 26-28, 2017.
- Anonby, Erik, Amos Hayes, J.P. Fiset and Robert Oikle. "A Multi-Dimensional Approach to Classification of Iran's Languages." In *Advances in Iranian Linguistics*. Amsterdam: John Benjamins, forthcoming.
- Anonby, Erik, Masoud Mohammadirad, and Jaffer Sheyholislami. "Kordestan Province in the *Atlas of the Languages of Iran*: Research Process, Language Distribution, and Language Classification." In *Advances in Kurdish Linguistics*, edited by Songül Gündoğdu, Ergin Öpengin, Geoffrey Haig and Erik Anonby. Bamberg, Germany: Otto-Friedrich-Universität Bamberg, forthcoming.
- Anonby, Erik, and Parisa Sabethemmatbadi. "Representing Complementary User Perspectives in a Language Atlas." In *Further Developments in the Theory and Practice of Cybercartography: International Dimensions and Language Mapping*, edited by D.R. Fraser Taylor, Kumiko Murasugi, and Erik Anonby. Amsterdam: Elsevier, forthcoming.
- Anonby, Erik, Parisa Sabethemmatbadi, and Amos Hayes. "Reconciling Contradictory Perspectives of Language Identity in the *Atlas of the Languages of Iran*." Paper presented at the Cognitive Science Colloquium Series, Carleton University, February 25, 2016.
- Anonby, Erik, and Laurentia Schreiber. "Reflections on Implementing the ALI (*Atlas of the Languages of Iran*) Questionnaire: Qashqai Turkic of Shiraz." Paper presented at the Workshop on Refining the ALI Questionnaire, Universität Bamberg, July 7-8, 2017.
- Anonby, Erik, and Mortaza Taheri-Ardali. "Linguistic Diversity and Language Contact in Chahar Mahal va Bakhtiari Province, Iran." Paper presented at the 7th International Conference on Iranian Linguistics (ICIL7), Lomonosov Moscow State University and the Institute of Linguistics of the Russian Academy of Sciences, Moscow, August 28-30, 2017.
- Anonby, Erik, and Mortaza Taheri-Ardali. "Bakhtiari." In *Language Contact and Language Change in West Asia*, edited by Geoffrey Khan and Geoffrey Haig, 445-480. Berlin: De Gruyter, 2019.
- Anonby, Erik, Mortaza Taheri-Ardali, and Adam Stone. "Linguistic Variation in Chahar Mahal va Bakhtiari Province, Iran." *Linguistic Typology* (forthcoming).
- Anonby, Erik, Mortaza Taheri-Ardali, et al. *Atlas of the Languages of Iran (ALI)*. Ottawa: Geomatics and Cartographic Research Centre, Carleton University, 2015-2018, <http://iranatlas.net>.
- Asadpour, Hiwa. "Atlas-e zabāni-ye rāyāne'i-ye Āzarbāyjān-e Gharbi" [A computerized language atlas of West Azerbaijan] (MA thesis, Tehran: Olum va Tahqiqāt University, 2011).

- Baziār, Hossein. "Tarh-e melli-ye guyeshshenāsi-ye Irān: Chehel rustā dar ostān-e Ilām" [The national plan of dialectology: forty villages in Ilam Province] (MA thesis, Tehran: Islamic Azad University, Central Tehran Branch, 1391/2012).
- Borjān, Habib. "Kashan ix. The Median Dialects of Kashan." *Encyclopedia Iranica* 16.1, 38–47. New York: Center for Iranian Studies, Columbia University, 2012, <http://www.iranicaonline.org/articles/kashan-ix-the-median-dialects-of-kashan>.
- Borjān, Habib. *Is There Continuity Between Persian and Caspian? Linguistic Relationships in the South-Central Alborz*. Ann Arbor: American Oriental Society, 2013.
- Borjān, Habib. "The Tabaroid Dialects of South-Central Alborz." *Acta Orientalia* 66, no. 4 (2013): 427–41.
- Bulut, Christiane. "Turkic Varieties in West Iran and Iraq: Representatives of a South Oghuz Dialect Group?" In *Turkic in Iran: Past and Present*, edited by Heidi Stein, 15–99. Wiesbaden: Harrassowitz, 2014.
- Bulut, Christiane. "Convergence and Variation in the Turkic Varieties of Iran: Examples from Qashqā'ī." In *Turks and Iranians: Interactions in Language and History*, edited by Éva Ágnes Csató, Lars Johanson, András Róna-Tas, and Bo Utas, 235–82. Wiesbaden: Harrassowitz, 2016.
- Chan, Eugene. *Numeral Systems of the World's Languages*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 1998–2018, <https://mpi-lingweb.shh.mpg.de/numeral>.
- Comrie, Bernard, and Norval Smith. "Lingua Descriptive Studies Questionnaire." Leipzig: Max Planck Institute for Evolutionary Anthropology, 1977, <https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/linguaQ.php>.
- Deyhim, Giti. *Atlas-e guyeshshenākhti-ye Qasrān-e Dākhel* [Dialectological atlas of Qasrān-e Dākhel]. Tehran: Sorush, 1388/2009.
- Doerfer, Gerhard. *Lexik und Sprachgeographie des Chaldasch* [Lexicon and linguistic geography of Khalaj]. Wiesbaden: Harrassowitz, 1987.
- Dryer, Matthew S., and Martin Haspelmath, eds. *World Atlas of Language Structures (WALS)*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013, <http://wals.info>.
- Ethnologue = Simons, Gary F., and Charles D. Fennig, eds. *Ethnologue: Languages of the World*. 21st ed. Dallas, Texas: SIL International, 2018, <http://www.ethnologue.com>.
- Farāhāni Jam, Farānak. "Tarh-e melli-ye Atlas-e Zabāni" [The national plan of the Language Atlas]. *Jām-e Jam Online*, 1388/2009, http://www.aftabir.com/articles/view/applied_sciences/social_science/c12c1240291037_language_p1.php.
- François, Alexandre. "Trees, Waves and Linkages: Models of Language Diversification." In *Routledge Handbook of Historical Linguistics*, edited by Claire Bowerman and Bethwyn Evans, 161–89. London: Routledge, 2014.
- GCRC (Geomatics and Cartographic Research Centre). *Nunaliit Atlas Framework*. Ottawa: GCRC, Carleton University, 2006–2018, <http://nunaliit.org>.
- GCRC (Geomatics and Cartographic Research Centre). "Developer Documentation: GCRC/nunaliit wiki-Github." Ottawa: GCRC, Carleton University, 2018, <http://github.com/GCRC/nunaliit/wiki>.
- Geiger, Wilhelm, and Ernst Kuhn, eds. *Grundriss der iranischen Philologie* [Outline of Iranian philology]. Strasbourg: Trübner, 1895–1904.
- Haig, Geoffrey, and Stefan Schnell. "Annotations Using GRAID (Grammatical Relations and Animacy in Discourse)." In *Multi-CAST*, edited by Geoffrey Haig and Stefan Schnell. Köln: Language Archive Cologne, 2014, https://lac.uni-koeln.de/corpora/Multi-CAST/_multicast_background/Annotations/HaigSchnell2014_GRAID-Manual7.pdf.
- Hammarström, Harald, Sebastian Bank, Robert Forkel, and Martin Haspelmath. *Glottolog*. Jena: Max Planck Institute for the Science of Human History, 2018, <http://glottolog.org>.
- Haspelmath, Martin. "How Hopeless is Genealogical Linguistics, and How Advanced is Areal Linguistics?" *Studies in Language* 28, no. 1 (2004): 209–23.
- Hāshemi, Ehsānollāh. "Shenāsāi-ye guyeshhā-ye Iran: Shahrestān-e Ardestān az ostan-e Esfahān" [The identification of Iran's dialects: Ardestan Shahrestan of Esfahan Province]. *Ayandeh* 16, no. 1 (1369/1990): 160–68.
- Haspelmath, Martin, and Uri Tadmor, eds. *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton de Gruyter, 2009.

- Hayes, Amos, and D.R. Fraser Taylor. "Developments in the Nunaliit Cybercartographic Atlas Framework." In *Further Developments in the Theory and Practice of Cybercartography: International Dimensions and Language Mapping*, edited by D.R. Fraser Taylor, Kumiko Murasugi, and Erik Anonby. Amsterdam: Elsevier, forthcoming.
- Hayes, Amos, Peter L. Pulsifer, and J.P. Fiset. "The Nunaliit Cybercartographic Atlas Framework." In *Developments in the Theory and Practice of Cybercartography: Applications and Indigenous Mapping*, edited by D.R. Fraser Taylor, 129–40. Amsterdam: Elsevier, 2014.
- Heggarty, Paul. "Prehistory Through Language and Archaeology." In *Routledge Handbook of Historical Linguistics*, edited by Claire Bowerman and Bethwyn Evans, 598–626. London: Routledge, 2014.
- Hourcade, Bertrand, et al. Irancarto. Paris and Tehran: CNRS / University of Tehran, 2012, <http://www.irancarto.cnrs.fr>.
- Iran Statistics Center. Public Census of Population and Settlement 2016. Tehran: Ministry of the Interior, Iran Statistics Center, 2011/2016, <https://www.amar.org.ir>.
- Izady, Michael Mehrdad. "Linguistic Composition of Iran." 2013, http://gulf2000.columbia.edu/images/maps/Iran_Languages_lg.png.
- Jahani, Carina, and Agnes Korn. "Balochi." In *The Iranian Languages*, edited by Gernot Windfuhr, 634–92. Abingdon, Oxon: Routledge, 2009.
- Jügel, Thomas. "On the Linguistic History of Kurdish." *Kurdish Studies* 2, no. 2 (2014): 123–42.
- Kalyan, Siva, and Alexandre François. "Freeing the Comparative Method from the Tree Model: A Framework for Historical Glottometry." In *Let's Talk About Trees: Tackling Problems in Representing Phylogenetic Relationships Among Languages*, edited by Ritsuko Kikusawa and Lawrence Reid. Osaka: National Museum of Ethnology, forthcoming.
- Kiā, Sādeq. *Vazhenāme-ye shast o haft guyesh-e Irāni* [Dictionary of sixty-seven Iranian dialects]. Tehran: IHCS Press, 1390/2011.
- Labex EFL (Empirical Foundations of Linguistics). "LC3: Languages, Dialects and Isoglosses of the Iran-Caucasus-Anatolia-Near-East Area" (project). Paris: Labex EFL, 2018, <http://www.labex-efl.com/wordpress/2018/02/26/lc3-languages-dialects-and-isoglosses-of-the-iran-caucasus-anatolia-near-east-area/>.
- Lecoq, Pierre. "Le classement des langues irano-aryennes occidentales." *Cahiers de Studia Iranica* 7: *Études irano-aryennes offertes à Gilbert Lazard*, 247–64. Paris: Association pour l'avancement des études iraniennes, 1989.
- Lecoq, Pierre. "Les dialectes du centre." In *Compendium Linguarum Iranicarum*, edited by Rüdiger Schmitt, 313–26. Wiesbaden: Reichert, 1989.
- Lecoq, Pierre. "Les dialectes du sud-ouest." In *Compendium Linguarum Iranicarum*, edited by Rüdiger Schmitt, 341–49. Wiesbaden: Reichert, 1989.
- Leskien, August. *Die Declination im Slawisch-Litauischen und Germanischen* [Declination in Slavic-Lithuanian and Germanic]. Leipzig: Hirzel, 1876.
- Mann, Oskar. *Die Mundarten der Lur Stämme im südwestlichen Persien* [The dialects of the Lur tribe in South-western Persia]. Berlin: Georg Reimer, 1910.
- Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath, and Magnus Huber, eds. *Atlas of Pidgin and Creole Language Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013, <http://apics-online.info>.
- Modarresi, Yahyā. Majmu'eh-ye maqālāt-e Mardomshenāsi, vizheh-ye: Shenāsāi-ye guyeshbhā-ye Irān 2 [Review of *Anthropology Proceedings, Special Issue: Identification of Iran's Dialects*, vol. 2]. *Farhang* 6 (1369/1990): 309–13.
- National Cartographic Center. National Database of Geographic Names of Iran. Tehran: Ministry of the Interior, National Cartographic Center, 2016, <http://gndb.ncc.org.ir>.
- Orywal, Erwin, ed. "Karte 10, Vorderer Orient: Sprachen und Dialekte" [Map 10, Middle East: languages and dialects]. *Tübinger Atlas des Vorderen Orients* (TAVO), series A, vol. 8. Wiesbaden: Reichert, 1988.
- Pāpoli Yazdi, Mohammad Hossein. "Naqsheh-ye parākandegi-ye zabānhā dar rustāhā-ye shomāl-e Khorāsān" [Language distribution map of northern villages in Khorasan]. *Majjaleh-ye tahqiqāt-e joghrafiāi* [Iranian Journal of Geographical Research] 10 (1367/1988): 21–42.

- Parmun, Yadollāh. *Tarh-e melli-ye Atlas-e Zabāni-ye Iran: Ketābcheh-ye jāme'-e tadvin-e Atlas (mostanad-sāzi, paygābdādeh-hā, naqshah)* [National plan of the Iran Language Atlas: a comprehensive booklet of atlas compilation (documentation, database, map)]. Tehran: Language and Dialect Research Center, Cultural Heritage, Handicrafts and Tourism Organization, 1386/2007.
- Parmun, Yadollāh, and Zohreh Zāre'i. *Atlas-e zabāni-ye Iran: Mabāni va raveshshenāsi. Gozāresh-e atlas-e zabāni-ye Esfahān* [Iran Language Atlas: fundamentals and methodology. A report on the Esfahan Language Atlas]. Esfahan: Cultural Heritage, Handicrafts and Tourism Organization of Esfahan Province, 1385/2006.
- Persian Academy (Farhangestān-e Zabān-hā va Guyesh-hā-ye Irāni). "Rāhnemā-ye gerdāvari-ye guyesh-hā" [Guide to the collection of dialects]. Tehran: Persian Academy, 2009, <http://www.persianacademy.ir/UserFiles/File/Gouyesh.pdf>.
- Purriāhi, Ma'sūd, and Farzād Foruzānfār. *Shenāsāyi-ye guyeshhā-ye Iran: Shābrestān-e Ardestān az Ostān-e Esfahān* [Identification of Iran's dialects: Ardestan Shahrestan of Esfahan Province]. Tehran: Iranian Anthropology Center, 1362/1983.
- Redard, Georges. "État des travaux et publication: quelques cartes onomasiologiques" [The state of work and publication: some onomasiological maps]. In *Arbeitspapiere* 13 [Working papers 13]: *Atlas des parlers iraniens—Atlas de l'Afghanistan* [Atlas of the Iranian languages—Atlas of Afghanistan], edited by Georges Redard et al., 7–19. Bern: Universität Bern, Institut für Sprachwissenschaft, 1974.
- Rostambeyk Tafreshi, Ātusā. "Atlas-e guyeshi va andāzegiri-ye fāselehā-ye guyeshi dar ostān-e Yazd" [Dialect Atlas and measurement of dialect distances in Yazd Province]. *Majjaleh-ye pazhubeshhā-ye zabānshenāsi-ye tatbiqi* [Iranian Journal of Comparative Linguistics Research] 5, no. 10 (1394/2016): 57–74.
- Rostambeyk Tafreshi, Ātusā. "Atlas-e guyeshi va andāzegiri-ye fāselehā-ye guyeshi dar ostān-e Hamedān" [Dialect Atlas and measurement of dialect distances in Hamadan Province]. *Majjaleh-ye jostārkhā-ye zabāni* [Iranian Journal of Language Related Research] 7, no. 1.29 (1395/2016): 59–80.
- Rural Network of Iran. *Roostanet*. Tehran: Rural Network of Iran, 2016, <http://map.roostanet.com>.
- Sabethematabadi, Parisa, Erik Anonby, and Jaffer Sheyholislami. "A Comparative Analysis of Language Classification Taxonomies in Kurdish, Tehrani, and Abade'i Language Communities." Paper presented at the 3rd International Conference on Kurdish Studies (ICKS3), University of Exeter, June 26–28, 2017.
- Salāmi, Abd-ol-Nabi. *Ganjineh-ye guyeshshenāsi-ye ostān-e Fārs* [Treasury of the dialectology of Fars Province] (7 vols.). Tehran: Farhangestān-e Zabān va Adab-e Fārsi [Academy of Persian Language and Literature], 1383–1393/2004–2014.
- Schleicher, August. "Die ersten Spaltungen des indogermanischen Urvolkes" [The first division of the early Indo-European people]. *Allgemeine Monatsschrift für Wissenschaft und Literatur* [General Monthly Review for Science and Literature] 1853: 786–7.
- Schmidt, Johannes. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen* [Relationships of the Indo-European languages]. Weimar: Böhlau, 1872.
- Schmitt, Rüdiger. *Compendium Linguarum Iranicarum* [Compendium of Iranian languages]. Wiesbaden: Reichert, 1989.
- Schreiber, Laurentia, Mortaza Taheri-Ardali, Geoffrey Haig, and Erik Anonby. "Contact-induced Change in Irano-Turkic Morphosyntax." Paper presented at the Conference on Language Contact and Language Change in Western Asia, Goethe University Frankfurt, March 10–12, 2017.
- Schreiber, Laurentia, Mortaza Taheri-Ardali, Geoffrey Haig, and Erik Anonby. "Language Contact as a Source of Morphosyntactic Variation in Three Varieties of Qashqay Turkic in South-West Iran." Paper presented at the 24th Linguistisches Internationales Promotionsprogramm (LIPP) Symposium, Ludwig-Maximilians-Universität Munich, June 21–23, 2017.
- Simons, Gary F., and Charles D. Fennig, eds. *Ethnologue: Languages of the World*. 21st ed. Dallas, Texas: SIL International, 2018, <http://www.ethnologue.com>.
- SJS = Sāzemān-e Jahād Sāzandegi. *Census of Rural Districts [dehestān] of Iran*. Tehran: Sāzemān-e Jahād Sāzandegi, 1986.

- Skjærvø, P. Oktor. "Iran vi. Iranian Languages and Scripts (2), Documentation." *Encyclopædia Iranica* 13, no. 4: 348–66. New York: Center for Iranian Studies, Columbia University, 2006, <http://www.iranicaonline.org/articles/iran-vi2-documentation>.
- Stilo, Don. "The Tati Language Group in the Sociolinguistic Context of Northwestern Iran and Transcaucasia." *Iranian Studies* 14, no. 3/4 (1981): 137–87.
- Stilo, Don. "Introduction to the Atlas of the Araxes-Iran (Outer Fertile Crescent) Linguistic Area." Paper presented at the School of Oriental and African Studies, London, 2007, http://www.hrhelp.org/events/seminars/abstracts/stilo_abstract.pdf.
- Stilo, Don. "Isfahān, xxi. Provincial Dialects." *Encyclopædia Iranica* 14, no. 1, 93–112. New York: Center for Iranian Studies, Columbia University, 2007, <http://www.iranicaonline.org/articles/isfahan-xxi-provincial-dialects>.
- Stilo, Don. "The Araxes-Iran Linguistic Area with a Concentration on Kurdophone Varieties." Paper presented at the 3rd International Conference on Kurdish Linguistics (ICKL3), University of Amsterdam, August 25–26, 2016.
- Stilo, Don. "Caspian and Tatic." In *Language Contact and Language Change in West Asia*, edited by Geoffrey Khan and Geoffrey Haig, 659–824. Berlin: De Gruyter, 2019.
- Swadesh, Morris. *The Origin and Diversification of Language*, edited by Joel Sherzer. Chicago, IL: Aldine, 1971.
- Taheri-Ardali, Mortaza. "A First Account of the Languages in Western Hormozgan." Poster presented at the 7th International Conference on Iranian Linguistics (ICIL7), Lomonosov Moscow State University and the Institute of Linguistics of the Russian Academy of Sciences, Moscow, August 28–30, 2017.
- Taheri-Ardali, Mortaza, and Erik Anonby. 2017. "Language Distribution in Chahar Mahal va Bakhtiari Province of Iran." Presented at 1st International Conference on Languages, Dialects and Linguistics (ICCILDL1), Pazhoheshgaran Andishmand Institute, Ahvaz, Iran, February 2–3, 2017.
- Windfuhr, Gernot. "Western Iranian Dialects." In *Compendium Linguarum Iranicarum* [Compendium of Iranian languages], edited by Rüdiger Schmitt, 294–5. Wiesbaden: Reichert, 1989.
- Windfuhr, Gernot. "Dialectology." *Encyclopædia Iranica* 7, no. 4, 362–70. New York: Center for Iranian Studies, Columbia University, 1995, <http://www.iranicaonline.org/articles/dialectology>.
- Windfuhr, Gernot. "Dialectology and Topics." In *The Iranian Languages*, edited by Gernot Windfuhr, 5–42. London: Routledge, 2009.
- Zhukovskij, Valentin A. *Materialy dlya izucheniya persidskikh narechij* [Materials for the study of Persian dialects]. St. Petersburg, 1888–1923.