

A Three-Dimensional Approach to Classification of Iran's Languages

Erik Anonby, Carleton University and Universität Bamberg

Language atlases typically approach language mapping from two directions: the distribution of language varieties – language families, languages and dialects – in geographic and social space, and an account of linguistic forms associated with these varieties.

But what varieties should be selected for a language atlas? How should they be grouped together and represented? An inventory of languages, along with their classification, is important in defining an atlas's outlook, scope, and visual representations.

Along with a complex language situation, fragmentary documentation, logistical problems, shortcomings in research design and unpublished results, persistent lack of agreement on the listing and classification of Iran's languages is a major obstacle to the production of a language atlas for the country since efforts first began in the 1950s (Anonby 2015).

In our current work on the *Atlas of the Languages of Iran* (<http://iranatlas.net/>), we are addressing fundamental issues in classification through an eclectic methodology that brings together existing work on documentation and language classification, theoretical insights on language identity and identification, and representational innovations. Five key elements of this area of research in the Atlas are as follows:

1. An ever-expanding inventory of all language varieties in Iran at all levels – language families, languages and dialects – based on the labels encountered in published sources and in fieldwork. At this point, we have compiled a list of over 500 distinct varieties, of which almost 400 belong to the Iranian family.

2. Construction of a working classification of all of these varieties using a traditional tree structure (<http://iranatlas.net/index.html?module=module.classification>). This portion of the research uses published classificatory work (e.g., Stilo 1981, Skjærvø 2006, Windfuhr 2009, Bulut 2014) as its starting point. In cases of divergences among scholars, we catalogue each point of view and the stated justification (if any) for a given classification. Finally, in the many cases where there are gaps and no classification is available for a given language variety or family, we have consulted senior figures in the field and developed our own assessments through field research.

3. Backgrounding of assessments of a given language variety as a language vs. dialect, dialect group, language family, etc. As shown in Anonby et al. (2016), there is no consensus about what constitutes a “language” and, by extension, what languages are spoken in Iran: whereas official administrative materials usually mention between four and seven languages (SJS 1986), scholars often cite dozens (TAVO 1988), and one source (Lewis et al. 2016) claims more than 70 distinct languages. Rather than attempting to defend a single, definitive list of languages, we concentrate on the nature and structure of relationships among all varieties.

4. Construction of a three-dimensional web model of language classification as an alternative (see <http://iranatlas.net/index.html?module=module.taxonomy.forceGraph#>) to a traditional two-dimensional tree. The three-dimensional geometry of this classification web, which is displayed using a “force graph” (Grandjean 2015), overcomes important limitations in the tree model (see Aikhenvald & Dixon 2001, François 2014) by allowing for complex parentage and facilitating the representation of more than one type of link between language varieties (see point #5).

5. Explicit differentiation and visualization of three types of links between language varieties: genealogical inheritance, structural similarity through contact, and ethnic identification.

- a. *Relation through genealogical inheritance.* The comparative method has as its goal the identification of genealogical inheritance, and we maintain this type of link as a central means of organizing language varieties. As in a traditional tree classification, language varieties can diversify over time and



result in multiple children. In the case of a mixed language (in the technical sense of the term, where a dual grammar exists) such as Kumzari (as described by van der Wal Anonby 2015), the three-dimensional geometry of the classification web allows for identification of genealogical parents from both contributing language families.

b. *Structural relation through contact.* In some cases, such as the Tabaroid (or Tatoid) group between Mazandarani and Tati (Borjian 2013, Stilo p.c. 2015), or the peripheries of the Lori language bloc (Anonby 2004/5), longstanding contact between languages – whether genealogically related or not – results in structurally intermediate varieties. A careful application of the comparative method can give insight into the most probable genealogical parent, and this link is maintained in our classification. However, the contact-induced “parentage” of the contact language is also represented.

c. *Relation through ethnic identification.* There are many situations where considerations of ethnicity are given a defining role in language classification, even among linguists. A well-known example is the conflation of Zazaki, Gorani, and Kurdish as a single branch within West Iranian (Windfuhr 2009, Kurdish Academy of Language 2014), even though the comparative evidence for such a grouping is lacking (Borjian p.c. 2013, Haig p.c. 2013). Northern Laki, on the other hand, is viewed by its speakers as a dialect of Lori, even though its basic structures pattern with Kurdish (Anonby 2004/5). Rather than rejecting this pervasive type of classification altogether, we have chosen to make the ethnicity-derived links explicit and to distinguish them clearly from genealogical classifications developed using the comparative method.

The overarching goal of this research is to develop a fuller understanding of the nature of Iran’s languages and their relationships to one another. Here, we have outlined current efforts to bring together existing perspectives on classification of Iran’s languages, and to complete a fully annotated and referenced three-dimensional classification of our own in the Atlas. Ultimately, comparison of classification-dependent language distribution maps with language data maps will lead to further refinements in the classification.

References

- Aikhenvald, Alexandra Y., and R.M.W. Dixon (eds.) 2001. *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford & New York: Oxford University Press.
- Anonby, Erik. 2004/5. Kurdish or Luri? Laki’s disputed identity in the Luristan province of Iran. *Kurdische Studien* 4/5:7-22.
- Anonby, Erik. 2015. Mapping Iran’s languages: Situation and prospects. ICIL6 (6th International Conference of Iranian Linguistics), Ilia State University, Tbilisi, Georgia, June 23-26, 2015.
- Anonby, Erik, Parisa Sabethemmatbadi and Amos Hayes. 2016. Reconciling contradictory perspectives of language identity in the Atlas of the Languages of Iran. CogSci Colloquium Series, Ottawa, February 25, 2016.
- Anonby, Erik et al. (eds.). 2016. *Atlas of the languages of Iran*. Ottawa: Geomatics and Cartographic Research Centre. Available at: <http://iranatlas.net>.
- Borjian, Habib. 2013. The Tabaroid dialects of south-central Alborz. *Acta Orientalia* 66.4, 427-441.
- Bulut, Christiane. 2014. Turkic varieties in west Iran and Iraq: Representatives of a South Oghuz dialect group? In Heidi Stein (ed.), *Turkic in Iran: Past and present*, 15-99. Wiesbaden: Harrassowitz.
- François, Alexandre. 2014. Trees, waves and linkages: Models of language diversification. In Bowern, Claire & Bethwyn Evans (eds.), *Routledge handbook of historical linguistics*, 161-189. London: Routledge.
- Grandjean, Martin. 2015. Introduction à la visualisation de données : l’analyse de réseau en histoire. *Geschichte und Informatik* 18/19, 109-128.
- Kurdish Academy of Language. 2014. *KAL Kurdish genealogy tree*. At: <http://kurdishacademy.org/?q=node/159>.
- Lewis, M. Paul et al. (eds.). 2016. *Ethnologue: Languages of the world*, 19th edition. Dallas, Texas: SIL International. Available at: <http://www.ethnologue.com>.
- TAVO = Orywal, E. (ed.). 1988. *Tübinger Atlas des Vorderen Orients (TAVO)*, A XIII. Wiesbaden: L. Reichert.
- Skjærvø, P. Oktor. 2006. Iran vi. Iranian languages and scripts: 2. Documentation. *Encyclopædia Iranica* 13.4, 348-366. New York: Center for Iranian Studies, Columbia University.
- SJS (Sāzemān-e Jahād Sāzandegi). 1986. *Census of rural districts [dehestān] of Iran*. Tehran: Sāzemān-e Jahād Sāzandegi.
- Stilo, Donald. 1981. The Tati language group in the sociolinguistic context of Northwestern Iran and Transcaucasia. *Iranian studies* 14.3/4, 137-187.
- Van der Wal Anonby, Christina. 2015. A grammar of Kumzari: A mixed Perso-Arabian language of Oman (Ph.D. dissertation). Leiden: Leiden University.
- Windfuhr, Gernot (ed.). 2009. *The Iranian languages*. Oxon & New York: Routledge.