

The Languages and Linguistics of Western Asia
WOL 6

The World of Linguistics

Editor

Hans Henrich Hock

Volume 6

De Gruyter Mouton

The Languages and Linguistics of Western Asia

An Areal Perspective

Edited by

Geoffrey Haig
Geoffrey Khan

De Gruyter Mouton

ISBN 978-3-11-042608-3

e-ISBN (PDF) 978-3-11-042168-2

e-ISBN (EPUB) 978-3-11-042174-3

Library of Congress Cataloging-in-Publication Data

Names: Haig, Geoffrey, editor. | Khan, Geoffrey, editor.

Title: The languages and linguistics of Western Asia : an areal perspective /
edited by Geoffrey Haig, Geoffrey Khan.

Description: Berlin ; Boston : De Gruyter Mouton, 2018. | Series: The world
of linguistics ; volume 6

Identifiers: LCCN 2018012560 | ISBN 9783110426083 (hardcover)

Subjects: LCSH: Middle East--Languages.

Classification: LCC P381.M53 .L36 2018 | DDC 409.5--dc23 LC record available at
<https://lccn.loc.gov/2018012560>

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed
bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2019 Walter de Gruyter GmbH & Co. KG, Berlin/Boston

Cover image: G. Haig

Typesetting: Dörlemann Satz GmbH & Co. KG, Lemförde

Printing and binding: CPI books GmbH, Leck

Acknowledgements

We have been very fortunate in having the assistance of various people in the preparation of this volume. In the early stages Nicholas Peterson (Bamberg) assisted in copy-editing, later Soner Ö. Barthoma (Cambridge) copy-edited several chapters, while the bulk of the final copy-editing and preparation of the indices was undertaken by Laurentia Schreiber (Bamberg). Nils Norman Schiborr (Bamberg) produced the maps. We gratefully acknowledge their valuable contribution to the project.

Many thanks also to the staff at Mouton de Gruyter for their guidance and efficiency during the long process of producing this volume: to Birgit Sievert for organisation and support in the early phases of the project, to Barbara Karlson for accompanying the editing process through to completion, and to Frauke Schafft for supervising the final production of the volume.

We would also like to thank all the scholars who participated in the project for their excellent contributions, and their enduring patience.

Finally, we would like to express our gratitude to the many anonymous speakers of the languages described in this volume, who provided much of the data on which the contributions are based. Without their interest and enthusiasm for their own languages, we could not have produced this volume.

Geoffrey Haig and Geoffrey Khan (editors)

November 2018

Table of contents

1. Introduction · Geoffrey Haig and Geoffrey Khan	1
2. Eastern Anatolia and northwestern Iran	31
2.1. Eastern Anatolia and northwestern Iran: overview · Geoffrey Khan	31
2.2. The Armenian dialects · Hrach Martirosyan	46
2.3. Northern Kurdish (Kurmanji) · Geoffrey Haig	106
2.4. The Arabic dialects of eastern Anatolia · Stephan Procházka	159
2.5. The Neo-Aramaic dialects of eastern Anatolia and northwestern Iran · Geoffrey Khan	190
3. Northern Iraq	237
3.1. Northern Iraq: overview · Geoffrey Khan and Geoffrey Haig	237
3.2. The Arabic dialects of northern Iraq · Stephan Procházka	243
3.3. The Iranian languages of northern Iraq · Geoffrey Haig	267
3.4. The Neo-Aramaic dialects of northern Iraq · Geoffrey Khan	305
3.5. Iraq-Turkic · Christiane Bulut	354
4. Western Iran	385
4.1. Western Iran: overview · Geoffrey Khan	385
4.2. The Turkic varieties of Iran · Christiane Bulut	398
4.3. Bakhtiari · Erik Anonby and Mortaza Taheri-Ardali	445
4.4. The Neo-Aramaic dialects of western Iran · Geoffrey Khan	481
4.5. Hawrāmī of western Iran · Parvin Mahmoudveysi and Denise Bailey	533
4.6. Persian · Ludwig Paul	569
4.7. Kumzari · Christina van der Wal Anonby	625
5. The Caspian region and south Azerbaijan: Caspian and Tatic · Donald L. Stilo	659

6. Caucasian rim and southern Black Sea coast	825
6.1. Caucasian rim and southern Black Sea coast: overview · Geoffrey Haig and Geoffrey Khan	825
6.2. Laz · René Lacroix	830
6.3. Ossetic · David Erschler	861
6.4. Romeyka · Laurentia Schreiber	892
7. Appendix: Basic lexical items in selected languages of Western Asia · Geoffrey Haig, Geoffrey Khan and Laurentia Schreiber	935
7.1. Background	935
7.2. Iranian languages	936
7.3. Semitic languages	941
7.4. Other language families	945
7.5. Notes on sources	947
Index of language names	953
Index of place names	961
Index of subjects	969

1. Introduction¹

Geoffrey Haig and Geoffrey Khan

1. Aims and scope of the volume

This volume contains a series of descriptions of around 20 languages (or language groups), spoken across a region that includes most of eastern Turkey, western Iran and northern Iraq (see Fig. 1 for the locations), together with overview articles of sub-regions, and an appendix with selected lexical items from these languages. The region that we refer to here loosely as “Western Asia” is not clearly demarcated, either politically or topographically, and requires a few words of explanation. Essentially it is linked to an ancient cultural core, namely the northern part of Mesopotamia, the upper catchment regions of the Tigris and Euphrates rivers. The book’s coverage goes beyond Mesopotamia in the narrower sense to include the surrounding mountainous regions of the Zagros in the southeast, and the elevated regions northwards and eastwards across the Anatolian plateau into the outliers of the Caucasus. These areas were all involved, at least economically and strategically, in the succession of empires that arose in and around ancient Mesopotamia, beginning with the Sumerians in the third millennium BCE and continuing down to the Ottoman Empire into the dawn of the twentieth century. The sense of a common *Kulturreum*, while difficult to delineate precisely, is reflected by the growing recognition of linguistic parallels shared among the region’s languages, some of which we take up below.

We have further divided the region into five sub-areas, each of which is treated in a section of the volume, with its own overview chapter. The individual language chapters are quite varied in composition, some focussing on a single language spoken in a geographically narrow location (e. g. Kumzari, chapter 4.7) while others cover geographically dispersed and internally diverse groups (e. g. Armenian, chapter 2.2, or Iran-Turkic, chapter 4.2). For this reason, it has not been possible to impose a single format on the presentations, but we have ensured that linguistic examples are glossed according to standard practice, that all contributions include a map, and where possible, a glossed text sample of natural language. In addition, an Appendix contains comparative lists of lexical items from a selection of the languages, plus some additional languages that were not treated in the main body of the book. The approximate locations of the languages and language groups covered in the volume is shown in Fig. 1, whereby the symbols are only intended

¹ We are very grateful to Christina van der Wal Anonby, Erik Anonby, René Lacroix and Ludwig Paul for comments and corrections on an earlier version of this introduction. We of course bear the responsibility for the remaining shortcomings.

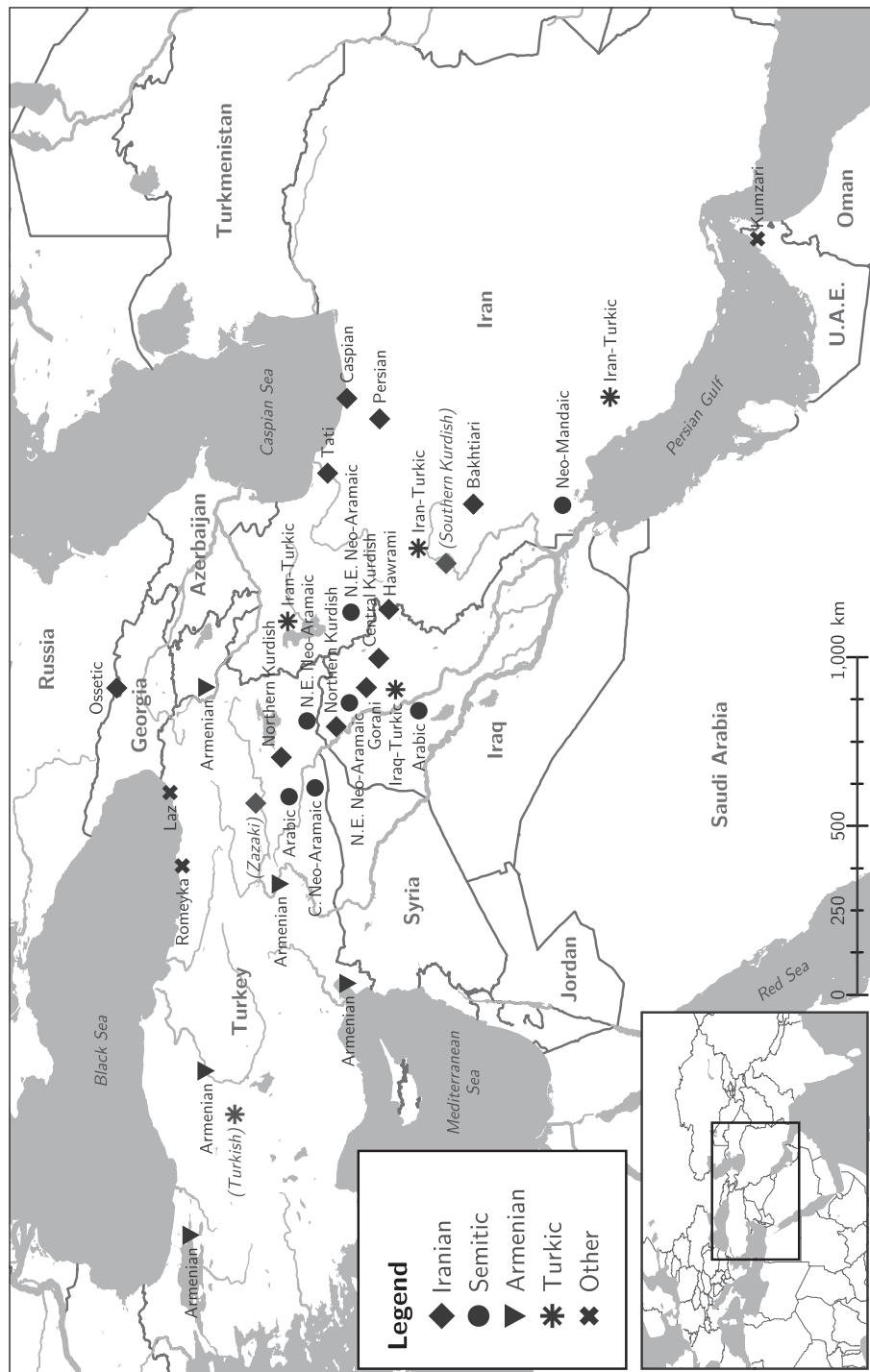


Figure 1: Approximate locations of main languages and language groups covered in this volume

to indicate the approximate centre of the respective language variety. Language names in brackets indicate important languages of the region, but which do not have a dedicated chapter in this volume.

The focus of the volume is on the minority languages of the region, for the most part under-described and endangered, while the respective state languages are generally outside the purview of the volume. This perspective is motivated by the recognition that taken together, the minority languages yield a much richer picture of the region's long-standing multilingual tapestry than the handful of currently dominant official languages (Arabic, Persian, Turkish). The latter have all been described exhaustively elsewhere, and have only achieved their current dominance very recently, hence give at best a distorted picture of the historical processes that engendered the region's linguistic composition (we have nevertheless included an overview chapter of Persian, chapter 4.6, because of its unique status as a hegemonic language across much of the region over at least two millennia).

The volume's scope is also restricted by practical considerations. For example, the Caucasus is largely excluded because it is treated within a companion volume in the same series. Zazaki, an Iranian language of Central Anatolia, is not included because there is already an extensive overview chapter on Zazaki in Windfuhr (2009). In other cases, gaps have arisen simply due to the unavailability of appropriate material or suitably qualified authors at the right time, or through considerations of space, or through various organisational setbacks that we faced over the six years of compilation. A particularly unfortunate omission is Domari, the language of the Doms of the Middle East (see Herin 2012 on Domari of Aleppo, and Matras 2012 for Domari of Jerusalem). Similarly, Southern Kurdish has not been treated in any depth (see Fattah 2000, and Belleli 2016). We can only trust that anyone familiar with the task of compiling a volume of this scope will appreciate that comprehensive coverage is an ideal that is seldom achieved. And while we are very much aware of the gaps in the current compilation, we are confident that the present volume provides the first reasonably representative survey of the linguistic diversity of the region.

We turn now to the major thematic focus of the volume, the areal perspective. By this we mean that the volume's structure reflects broad areal divisions, rather than genetic groupings of the languages concerned. In keeping with the general approach, we have eschewed in-depth historical treatment of the language families; all the major families and branches have been dealt with extensively in recent handbooks (see for example Windfuhr 2009 on Iranian, Johanson and Csató 2006 on Turkic, and Weninger et al. 2011 on Semitic). With regard to the ancient language contacts of the region, we refer to Butts (2015) on Semitic, and Hasselbach-Andee (in prep.).² The individual chapters of the present volume primarily

² We are grateful to Ilya Yakubovich for drawing our attention to this latter volume in preparation.

focus on a synchronic description of the respective language or language group, but with explicit reference to contact issues. In the overview chapters to each sub-region, we provide a short synopsis of the most striking contact phenomena. Interest in areal contact across the region has steadily increased in the last years (see for example the joint Frankfurt/Cambridge initiative on language contact and language change in Western Asia,³ and the recent study on relative clauses from an areal perspective by Gandon 2016). Because the present volume combines expertise from diverse philological traditions, it will hopefully become a valuable one-stop resource for future research in this rapidly growing field.

2. Western Asia from an areal typological perspective

Within language typology, the relevance of large-scale areal units in shaping the global distribution of typological features is increasingly emphasized. The region under consideration here has not, however, generally been recognized to date. Dryer's (1992) proposal has been influential, and recognizes the following six macro-areas (see Hammarström and Donohue 2014 for critical discussion):

- Africa
- Australia-New Guinea
- Eurasia
- North America
- Southeast Asia and Oceania
- South America

On this coarse-grained view, our “Western Asia” would be assigned to “Eurasia”, which also includes e. g. Mongolian, or Mandarin. However, it is evident that the Semitic languages represented in Western Asia (e. g. Neo-Aramaic) are the north-easterly outliers of the Afro-Asiatic language family, and are thus historically rooted in Africa (see in this respect Güldemann's notion of “Afrabia”, Güldemann, 2018.). More recently, Bickel et al. (2017) propose a finer-grained division, which recognizes a macro-area “Greater Mesopotamia”. This more closely corresponds to our Western Asia, though in Bickel et al. (2017), it also includes the entire Arabian peninsula. But ignoring for a moment the issue of the precise boundaries,⁴ there are in fact reasons to consider Western Asia a special kind of linguistic area, which has hitherto not been afforded much attention.

³ See e. g. <http://titus.fkidg1.uni-frankfurt.de/click/fourthAct.html>.

⁴ It is worth pointing out that many of the regions traditionally considered to be linguistic areas (or *Sprachbünde*) lack clearly identifiable borders (e. g. “the Balkans”), and are defined in different ways by different scholars.

The linguistic diversity of the region is, within the broader Eurasian context, relatively high: four distinct language families⁵ are represented (Turkic, Semitic, Indo-European, Kartvelian), and within Indo-European, four branches (Armenian, Iranian, Indic, and Hellenic). But with the exception of Armenian, none of these groupings are indigenous to the region: all the languages represent outliers of larger groups, and have close relatives outside the immediate region. The high level of linguistic diversity has thus arisen secondarily, as it were, through the region's location at the intersection of several large genetic groupings. This is in stark contrast to, for example, the linguistic diversity of the Caucasus (Nichols' 1992 canonical example of a "residual zone"), stemming from three indigenous language families, all of which lack (proven) relatives outside the region itself. In a number of publications Stilo (e.g. 2005, 2012) looks at the region centred on the catchment of the Araxes river, considering it to be an "intersection zone". Among the most striking phenomena is the areal distribution of adpositional typologies, which shift from Turkic-type postpositional, across Iranian mixed typologies, to Semitic prepositional. Haig (2001, 2017) considers eastern Anatolia as a "transition zone", with gradual areal shifts across a number of morphosyntactic features, but also its own set of defining features. Among them is the common Anatolian OVX word order pattern (where "X" refers to various kinds of non-direct object arguments, in particular goals, addressees, and recipients, see §3.4.3 below). Like the mixed adpositional typologies of Iranian, the OVX word order can be considered a compromise strategy between OV and VO word orders. These kinds of typologically unusual patterns illustrate that a transition zone is not merely the sum of the participant language families, but may develop its own profile with typologically rare structures – rare simply because transition zones themselves emerge only under very specific, hence globally very rarely occurring, conditions. If large-scale macro-areas are relevant to understanding the distribution of structural features, then we suggest that the transitional zones that lie between them also merit close scrutiny, precisely because they involve the comparatively rare areal contiguity of different language types, yielding configurations that are unlikely to arise elsewhere.

3. Areal patterns in phonology and morphosyntax across Western Asia

This section discusses a selection of features from phonology and morphosyntax that we consider are evidence for the role of language contact in shaping the structures of the languages of the region. In some cases, the phenomena are regionally quite restricted, while others span a large cross-section. Though the language

⁵ We are ignoring some of the more recent arrivals to the region, such as communities of Chechen speakers in Anatolia.

sample we draw on is by no means comprehensive, it is nevertheless possible to draw a few more general conclusions regarding the range and constraints of contact-induced change in the languages of the region.

3.1. Language contact: general considerations

Two or more languages spoken in the same region may exhibit structural similarities through (i) mere chance, (ii) inheritance from a common ancestor, (iii) contact influence, or (iv) a combination of (i)–(iii). We are primarily interested in (iii), but assigning similarities to contact influence requires caution, and should only be undertaken after due consideration of other possible causes. First, we need to bear in mind the histories of related languages spoken outside the region (to the extent that they are known) in order to formulate a benchmark of comparison against which the developments in the languages under consideration can be assessed.

As an example of how the facts from related languages can be brought to bear in assessing contact influence, let us consider the example of North-Eastern Neo-Aramaic (NENA) and Central Neo-Aramaic (CNA) in the region (Khan, this volume, chapters 2.5, 3.4, and 4.4). These languages have developed ergative (or split-intransitive) alignments with their perfective verb forms, matching a broadly similar alignment profile in the neighbouring varieties of Kurdish. Alignment systems have often been considered relatively immune to areal influence (Nichols 1992: 181), so in principle, the similarities between Neo-Aramaic and Kurdish could have occurred independently, and we need solid arguments to make a case for contact influence. In this case, evidence from the related languages, but outside the immediate contact region, is extremely relevant. With regard to Semitic, the Neo-Aramaic development of ergativity is unique within the Semitic language family as a whole. It is absent in Western Neo-Aramaic in Syria, a close relative of NENA and CNA outside of the area. This fact greatly strengthens the argument in favour of assuming contact influence: if it were a solely language-internal development, we might have expected traces of it within related languages outside the geographical contact zone, but this does not seem to be the case. Furthermore, contact between Kurdish and the Neo-Aramaic speech communities of the region has been intense and continuous for many centuries, and there is widespread evidence of borrowing in other parts of the grammar (Khan 2007), lending further credence to the assumption of structural convergence. Considering the histories of the respective language families is also important in order to establish the direction of influence. We know that some kind of tense-sensitive ergativity is widely attested in Iranian languages spoken beyond the contact zone with Neo-Aramaic (e. g. Pashto, or eastern varieties of Balochi, or Tatic, Haig 2008). Thus the simplest assumption is that ergativity developed in Kurdish through inheritance from its Iranian ancestors, while its occurrence in Neo-Aramaic is (at least in part)

due to influence from an Iranian contact language, e. g. precursors of Kurdish, or Gorani.

A second factor to consider in assessing contact influence is typological probability. Consider for example the fact that most of the languages of the region have pre-verbal negation in simple clauses, with the exception of Turkic, and Kumzari (van der Wal Anonby, this volume, chapter 4.7). In Turkic, negation is via a verbal suffix, and Turkish is generally suffixal, so post-verbal negation is structurally pre-figured in the language. We might then be inclined to ignore Turkic, and consider the otherwise prevalent pre-verbal negation to be an areal trait. But in fact, pre-verbal negation is the most widespread kind of negation across the world's languages (Dryer 1988), so the case for areal influence in Western Asia is weak, and would need to be reinforced through, e. g., evidence of non-random similarities in the forms involved, if it were to be maintained. The presence of post-verbal negation in Kumzari, on the other hand, obviously requires additional assumptions: it does not conform to the typologically more frequent pattern, and is distinct from Kumzari's Iranian relatives, which have consistent pre-verbal negation. Thus contact influence from neighbouring South Arabian languages appears to be a very plausible causal factor in the emergence of post-verbal placement of negation markers in Kumzari, as noted in van der Wal Anonby (chapter 4.7, §7).

Several suggestions have been made for the relative ease of "borrowability" among different categories of morphosyntax (see Matras 2007 on hierarchies of borrowability, and see below for a proposal in this direction regarding syntax). While it is often stated that, e. g. derivation is more easily borrowed than inflection, or clitics are more easily borrowed than affixes, it is useful to identify the more general principles at work (Haig 2014a). In general, ease of borrowability is characteristic for items that (i) are perceptually salient (e. g. syllabic, as opposed to segmental, concatenative as opposed to non-concatenative); (ii) exhibit predictable and transparent form-meaning correspondences, with tangible semantics; (iii) have a greater degree of positional freedom (e. g. are phonologically independent items, with some versatility in positioning, as opposed to affixes locked into morphological templates).

From these it follows fairly naturally that words are more likely to be borrowed than clitics, and clitics more likely than affixes. The least borrowable items are irregular, non-concatenative exponents of an abstract grammatical category, such as the suppletive forms for the comparative grade of the adjectives *good* and *bad*, or irregular past-tense formations involving Ablaut (*give, gave* etc.) in English. Particularly striking evidence for the difficulty of borrowing non-concatenative (i) and unpredictable (iii) morphology can be seen in the fact that Semitic root-and-pattern morphology (e. g. broken plurals, or verb stem alternations in Arabic) is very rarely transferred to native lexical items in contact languages (but see van der Wal Anonby, this volume, chapter 4.7 on the incorporation of Arabic morphology into Kumzari and Paul, this volume, chapter 4.6, §4.4, for examples of

broken Arabic plurals borrowed from Arabic). At the other end of the spectrum, a derivational suffix such as Turkic *-či* (with various vowel-harmonic variants), meaning ‘occupation associated with the base’ (e.g. *saat-či* ‘watchmaker’) turns out to be readily borrowable, and occurs in most of the languages in the region (e.g. in the Muş dialect of Northern Kurdish *mesî-čî* ‘fisherman’, from Kurdish *masî* ‘fish’, example from Songül Gündoğdu p.c.). The suffix *-či* is perceptually salient, affixal, and regular in form and meaning.

Finally, a useful distinction is drawn by Matras (2007) between **matter** borrowing, and **pattern** borrowing. The former involves the borrowing (or copying) of items together with the phonological substance of the donor language. A clear case of matter borrowing is found in Turkic varieties spoken in northern Iraq, which have borrowed a comparative suffix *-tär* from neighbouring Iranian languages (Bulut, chapter 4.2, §2.3.1.4). Pattern borrowing involves the borrowing of, e.g. the structural organization of paradigms, or the relative ordering of morphemes in a word, or specific principles of form-meaning correspondence, but without borrowing of actual phonological material. An example of pattern borrowing is *m*-initial reduplication, where a word is reduplicated, but the initial segment replaced by [m]: Turkish *para mara* ‘money and stuff’, Laz *dadzi madzi* ‘thorns and stuff’ (Lacroix, chapter 6.2, §3.4). This is widespread across the region (see e.g. Khan 2016, vol. 2: 99 for NENA, Haig 2001 for other languages of Anatolia). Of course matter and pattern borrowing are not mutually exclusively. Often, pattern borrowing is accompanied by phonological similarity of the items concerned, as in the *m*-segment in the reduplication example just mentioned.

3.2. Phonology

A number of languages of the region have increased their consonantal inventory by acquiring consonants from other languages by areal diffusion. A particularly clear case of this is the spread of glottalized consonants, realized as unaspirated stops or ejectives, across the northern sector of the region, including Armenian dialects, Northern Kurdish dialects of eastern Anatolia and northwestern Iran, the Neo-Aramaic dialects belonging to the NENA group spoken in this Northern Kurdish area, and some languages of the Caucasian rim, such as Laz and Ossetic. In Laz glottalization is a feature inherited from Kartvelian. In Armenian it has considerable historical depth, as shown by the fact that it occurs in Classical Armenian, and is possibly an inheritance from earlier Indo-European (Martirosyan, chapter 2.2, §4.1), but it may well have been reinforced by the fact that it is an areal feature. In Kurdish, Neo-Aramaic and Ossetic it is not an inherited feature and has entered these languages by areal diffusion. This is clearly demonstrable in the case of Kurdish and Neo-Aramaic, since dialects of these languages spoken further south in northern Iraq and western Iran do not have glottalized consonants. The area of glottalized consonants extends further north into the Caucasus region, where it is

found in Nakh-Daghestanian languages (e. g. Archi, Lak), Northwestern Caucasian languages (e. g. Kabardian) and South Caucasian languages such as Georgian, which are not covered by this volume (see Grawunder 2017 on the areal distribution of phonological features in the Caucasus). The nature of glottalization varies across the area with regard to its degree of strength. In the languages of the Caucasus it is generally realized as glottal ejection. In languages on the Caucasus rim and in Armenian, Kurdish and Neo-Aramaic, which are described in this volume, the glottalization is weaker and is typically realized as the lack of aspiration without clear ejection. It is noteworthy, however, that dialects of Armenian and Neo-Aramaic that are spoken by migrant communities in Georgia pronounce glottalized consonants as ejectives (Chirikba 2008: 44–45; Khan 2016, vol. 1: 93–95), which is a further demonstration of the areal nature of the feature. This also shows that intensity of contact in the apparent epicentre of the areal feature in the Caucasus can bring about greater degrees of areal convergence.

The fact that languages and dialects outside of the Caucasus typically realize the glottalized consonants as unaspirated stops rather than ejectives gives some insight into how these consonants were diffused. An insightful case study in this respect is the development of the unaspirated stop phonemes in the Neo-Aramaic dialects. In Neo-Aramaic dialects that have glottalized consonants fully established in their sound inventories, such as the Christian NENA dialect of Urmi, these consonants occur predominantly in native Aramaic words. In fact it is difficult to identify any loanword in the Christian Urmi dialect that has a glottalized consonant in the source language. The glottalized phonemes have rather developed by internal processes. One such process is the reanalysis of the deaspirated allophones of stops in fricative—stop clusters as unaspirated phonemes, e. g. /xp/ [x̪] > /x̪/.⁶ Another is the reanalysis of the sequence of an aspirated stop and laryngeal stop as an unaspirated phoneme, e. g. /t/ + /ʔ/ > /ʃ/. A further process arises from the fact that in Christian Urmi Neo-Aramaic and neighbouring dialects, the original pharyngealization of pharyngealized consonant segments has developed into a suprasegmental feature that takes a whole word as its domain. In such dialects an original pharyngealized stop /t/ [t̪] becomes reanalyzed as an unaspirated stop segment /t/ within the suprasegmental domain of pharyngealization (for details see Khan 2016, vol. 1: 92–131). The new glottalized phonemes formed by these internal processes have then spread to other contexts and serve various functional purposes, such as diminutive sound symbolism, e. g. *t̪olpa* ‘eyelash’ (< **t̪olpa*) (Khan 2016, vol. 1: 180). The basic process, then, is that of phonemicization of already existing sound patterns within Neo-Aramaic that match or approximate to those of a contact language. This would have taken place by the perception of an equivalence between these matching or similar sounds and the glottalized phonemes of the contact language. Such innovative unaspirated stop phonemes

⁶ In the transcription unaspirated stops are distinguished by diacritics thus: *ḥ*, *ṭ*, *k̪*.

could then undergo further phonetic change by converging with the phonetic features of ejectives in languages with such ejective glottalized consonants by a “perceptual magnet effect”, as Blevins (2017) puts it. As mentioned, dialects of Northern Kurdish (*Kurmanjî*) spoken in close proximity to Armenian also possess unaspirated voiceless stops (in some dialects with ejective characteristics) in their native lexicon. These elements are not directly inherited from Iranian, but likewise cannot be easily explained in terms of a spread from loanwords; the number of clearly identifiable Armenian loans in the everyday lexicon of *Kurmanjî* is very small. We see, then, that the presence of perceptually salient segments in a contact language may act as attractors in guiding the direction of internal changes.

Conversely, languages that have inherited glottalized consonants may undergo weakening or reduction of this feature, if they are geographically isolated from languages with it. This is the case in Laz (Kartvelian), where the inherited glottalized consonants are weakened in comparison to related Georgian (Lacroix, chapter 6.2, §6.1), presumably due to the isolation of Laz from its Kartvelian relatives, and the influence of neighbouring Turkish. Nevertheless, Lacroix still posits glottalized consonants for Laz. Full loss of glottalization is apparently a rare phenomenon in the Western Asian context. Glottalized consonants appear to be thus (a) fairly resistant to loss, and (b) prone to spread in contact situations.

Pharyngeal and pharyngealized consonants, which are an inherited feature of the sound inventories of Semitic languages, have spread by diffusion to some non-Semitic languages of the region. The process of areal diffusion is not so clear-cut as that of glottalized consonants. Several non-Semitic languages exhibit pharyngeals in loanwords from Arabic. Some Iranian languages, such as Kurdish and the languages of Kordestan and Kermanshah provinces, have developed innovative pharyngeals in native lexical items, in particular the unvoiced pharyngeal, which, for example, is commonly found in these languages in the numeral ‘seven’ (Northern Kurdish, Hawrami *haft*, Central Kurdish *hawt*). The Semitic inherited pharyngeals, unvoiced [ħ] and voiced [ʕ], are found in the Arabic dialects of the region and in the Central Neo-Aramaic dialects of southeastern Turkey. In the NENA dialect group situated to the east of the Tigris, however, the inherited pharyngeals have been mostly lost, especially in many of the dialects in the eastern sector of the NENA area in the Hakkari mountains of Turkey and northwestern Iran. The preservation of pharyngeals in Central Neo-Aramaic and their partial preservation in some NENA dialects is likely to be conditioned by extensive contact with Arabic. This would, therefore, be a case of language contact preserving sound patterns. When pharyngeals occur in the aforementioned subgroup of NENA dialects these are predominantly inherited pharyngeals in the environment of pharyngealized consonants, e. g. Qaraqosh *ta'ən* [tˤɑ:ən] ‘he bears’, pharyngeals arising from the debuccalization of pharyngealized consonants, e. g. Jewish Sanandaj *'ahra* < **'aˤtˤrˤaˤ* < **'atrā* ‘town’, or innovative pharyngeals resulting from strengthening of laryngeals for functional purposes. The latter category includes cases

of the strengthening of laryngeals to preserve morphological distinctions, e. g. Qaraqosh 3ms possessive suffix *-əħ* < *-əħ to prevent the loss of final *-h*, which would have rendered the suffix homophonous with the noun plural suffix *-ə*, e. g. *tor-əħ* (< **tor-əħ* ox-POSS.3MS) ‘his ox’ vs. *tor-ə* (ox-PL) ‘oxen’. The emergence of pharyngeals in some Iranian languages possibly also has the functional motivation of strengthening distinctions in paradigms, e. g. Northern Kurdish, Dohuk: *ħaft* (< **ħaft*) ‘seven’ to distinguish it clearly from the similar sounding numeral *ħašt* ‘eight’.⁷

Glottalized consonants and pharyngeals are cross-linguistically relatively rare consonants (Maddieson 2013a, 2013b) and their diffusion has resulted in the enrichment of the consonantal inventories of the languages of the region. A factor that may have facilitated their “magnet” effect is their salience (Blevins 2017). This would have conditioned not only their tendency to spread but also their resistance to loss. It is relevant in this respect that the unvoiced pharyngeal [ħ] diffused more readily than the voiced pharyngeal [ʕ], no doubt because the unvoiced member of the pair is the more salient due to its higher pitch.

Contact has resulted in loss of consonants of a number of languages of the area. A clear case of this is the loss of inherited interdental consonants of some Arabic and Neo-Aramaic dialects. This has brought the phoneme inventories of these dialects closer to those of the non-Semitic contact languages, all of which lack interdentals. The most common process involves merger of the interdentals with other consonants that have a direct match in the inventories of the contact languages. These are typically stops or sibilants. In many cases this merger is symmetrical, e. g. all the interdentals merge with stops (Arabic θ, ð, ḫ > t, d, d̪, NENA θ, ð > t, d) or sibilants (Arabic θ, ð, ḫ > s, z, z̪, NENA θ, ð > s, z). Sometimes, however, there is asymmetry. This is found in a number of NENA dialects. In such cases the outcome results in the reflex of the unvoiced interdental being weaker than that of the voiced interdental. In some dialects the unvoiced interdental is preserved while the voiced one shifts to a stop, e. g. Christian Ankawa *xaθa* ‘sister’, *'ida* (< **'iða*) ‘hand’, the unvoiced becomes a sibilant, while the voiced becomes stop, e. g. Jewish Nerwa *xasa* (< **xaθa*) ‘sister’, *'ida* (< **'iða*) ‘hand’, or the unvoiced undergoes debuccalization while the voiced becomes a stop, e. g. Jilu ⁺*xa* ‘sister’ (< **xaθa*),⁸ *'ida* (< **'iða*) ‘hand’. Some unusual outcomes include the shift of the Arabic interdentals θ, ð, ḫ to labio-dental fricatives f, v, y in some Arabic dialects (Procházka, chapter 2.4, §1.1) and the shift in some Jewish Neo-Aramaic dialects of both the interdentals θ and ð to the lateral l (Khan, chapter 2.5, §5.1, chapter 3.4, §5.1, chapter 4.4, §5.1.1). The Neo-Aramaic dialects that have this latter outcome were spoken in areas where in neighbouring Iranian and Turkic languages a /d/

⁷ Data from Matras, Yaron et al. 2016. *The Dialects of Kurdish*. Web resource, University of Manchester. <http://kurdisch.humanities.manchester.ac.uk/> (accessed 17 March 2018).

⁸ The superscribed symbol ⁺ denotes suprasegmental pharyngealization.

following a vowel or sonorant undergoes lenition, known as “Zagros *d*”, resulting in it being realized as an approximant or as sonorant (Haig, chapter 3.3, §3.1.1; Mahmoudveysi and Bailey, chapter 4.5, §3.1; Anonby and Taheri-Ardali, chapter 4.3, §2.1). There is evidence that in the Neo-Aramaic dialects concerned the two interdentals first shifted to the voiced stop **d* before finally becoming a lateral sonorant /l/. The sonorant /l/ can then be regarded as lenition of the **d*. Such lenition, therefore, is likely to be due to the “perceptual magnet effect” of the weakened Zagros *d*, whereby Neo-Aramaic speakers match this perceptually with the sonorant /l/ in their existing sound inventory. We may conclude that interdentals are prone to loss and do not spread by contact. This can be correlated with their lack of salience (Maddieson 2013b).

Shifts in the realization of consonants without necessary loss to consonantal inventories also result from areal diffusion. A case of this is the palatalization of the dorsal stops /k/ and /g/, which has occurred in many languages of the region. This is found, for example, in several Turkic varieties, in particular in the Turkic spoken in the north of the region in the Azerbaijaniyan enclave Nakhichevan, in some regions of Georgia, around Tabriz and Urmia in Iran, and in northern Iraq (Bulut, chapter 4.2, §2.1.2.4). Similar palatalization of these consonants is found in some NENA dialects in Iraq and northwestern Iran (Khan, chapter 3.4, §5.1, Khan, chapter 3.5, §5.1), in many Iranian languages, including those of the Caspian region (Stilo, chapter 5, §3.1), various Kurdish dialects of Anatolia, Iraq and northwestern Iran (chapter 2.1, §1.5, Haig, chapter 3.3, §3.1.1), Ossetic (Erschler, chapter 6.3, §2.1) and Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §2.1). It is also found in Romeyka (Schreiber, chapter 4.4, §2.1.2). The degrees of palatalization vary, ranging from stops with palatal offglides to a shift to the affricates [tʃ] and [dʒ]. In some languages the palatalization is conditioned by adjacent high vowels (e.g. in Ossetic and some languages in the Caspian region), whereas in others it is not conditioned by the environment. Palatalization appears to have been diffused as an areal feature. This is shown, for example, by the fact that NENA dialects in Iraq and Iran only exhibit this process in areas where it is found in the contact languages Turkic and Kurdish (chapter 3.1, §2). The varying degrees of palatalization can be correlated in some cases with different degrees of intensity of contact. This is shown, for example, by the case of the NENA dialect spoken by Christians in the plain of Urmia in northwestern Iran. In varieties of this dialect spoken in the southern areas of the plain palatalization results in affrication whereas in the varieties spoken in the north of the plain the dorsal stops are not fully affricated. This can be correlated with the fact that in the south the NENA speakers frequently code-switched between NENA and Azeri Turkic, whereas this was rare in the north. Azeri Turkic exhibits affrication of dorsals. Moreover, the Mukri Kurdish dialect spoken in the south of the Urmia plain has affrication of dorsals, but this is not found in the Kurdish spoken in the north of the plain. All this meant that the speakers of NENA in the south had greater contact with affrica-

tion of dorsals in the contact languages than speakers in the north (for details see Khan 2016: 111).

In the vowel systems of the languages of the region a clear case of areal diffusion is the fronting of the high back vowel [u:] to [y:] (transcribed *ii* below). Front rounded vowels are generally considered marked in comparison to back rounded vowels. The only languages of the region to have historically inherited front rounded vowels are the Turkic languages. Interestingly, among the languages treated in this volume which have undergone heavy contact influence from Turkic, none are analysed as having front rounded vowels in their vowel system. These languages include Romeyka (Schreiber, chapter 6.4), Anatolian dialects of Arabic (Procházka, chapter 2.4, see also Jastrow 2011), Laz (Lacroix, chapter 6.2), Northern Kurdish of Anatolia (Haig, chapter 2.3). This is particularly noteworthy for the Arabic dialects of Siirt, Kozluk and Sason, discussed in Jastrow (2011). Although the vowel systems (and some parts of morphosyntax) have been extensively restructured under Turkish influence, they apparently do not include front rounded vowels.

Front rounded vowels are found in the Kurdish dialects in southeastern Anatolia (Şemzînan) and northwestern Iran, and in the Bahdini Kurdish dialect of northern Iraq. This results in a pull-chain effect whereby [o:] is raised to [u:] to fill the gap left by the fronting of original [u:] (Haig, chapter 3.3, §4.1.2). The same processes are found in NENA dialects that were in contact with these specific Kurdish dialects, e. g. *‘Mawana: xabūša < *habbūšā ‘apple’, ruqe < *rōqē ‘spittle’*. The emergence of this areal feature appears to have come about through processes of fronting rather than under the influence of Turkic. This can result in partial fronting outcomes. In the Christian Urmi dialect of NENA, for example, long *ii* is not fully fronted but shifts to a diphthong with a palatal offglide: *xabuyša ‘apple’* (Khan, chapter 2.5, §5.5.2). In this latter case the perceptual magnet of the Kurdish front rounded vowel has resulted in only a partial convergence. In the Christian NENA dialect of Salamas the offglide is pushed up rather than forward, and is realized as a velar fricative *x*, e. g. *xibuxša < *habbūšā ‘apple’* (Khan, chapter 2.5, §5.5.2). This may be a strategy to create a sound pattern that is more saliently distinct from *u* than *uy*, suggesting that it was a later internal development in this Neo-Aramaic after a diphthong *uy* had arisen by contact with the Kurdish fronted rounded vowel. Fronting may be conditioned by assimilation processes at morpheme boundaries, as in Gorani dialects of western Iran: [ø] occurs in some verb forms, e. g. [mønim] ([<mæ-win-im < *mæ-bin-im], ‘I see’, Mahmoudveysi et al. 2012: 11). An exceptional case of the emergence of front rounded vowels by diffusion from Turkish is Asia Minor Greek of Cappadoccia (not treated in this volume, see Dawkins 1916), which is frequently cited as a language exhibiting exceptionally high levels of influence from Turkish. Turkic languages that have been heavily influenced by Iranian languages, on the other hand, have often lost at least some of their front rounded vowels. The most extensive loss is found in Turkic varieties in

Iran and the southernmost varieties of Iraq Turkic, Mandali and Khanaqin. Bulut (chapter 3.5, §2.1.2.3, chapter 4.2, §2.1.2.1) gives examples of the shift of the front rounded vowels /ü/ and /ö/ to their unrounded counterparts /i/ and /e/ (or [e]), such as /sit/ < *sit* ‘milk’, and [e:z] < *öz* ‘self, own’, or [dek-] < *dök-* ‘to pour’.

The backing of the vowel *a* is a feature that has diffused across several languages in the Anatolian region. In Central Neo-Aramaic spoken in the area of Tûr ‘Abdîn all historical long *a* vowels have the quality of *o*, e. g. Turopyo: *hmoro* ‘ass’ (< **hmārā*). This is a feature also of modern Western Aramaic dialects spoken in Syria (Arnold 2011) and can be traced to a considerable time depth in Aramaic in the Levant and other ancient Semitic languages of this region, such as Phoenician and Hebrew. It is likely to have diffused to Central Neo-Aramaic dialects of Turkey from the Levant at an early period. NENA dialects that were spoken in Turkey adjacent to Central Neo-Aramaic exhibit some degree of convergence with this vowel typology. In some NENA dialects of the Bohtan area long *a* in stressed syllables shifts to *o*, e. g. Ruma *xmóra* ‘ass’ (< **hmārā*). This may be explained as the perceptual coupling of long *ā* in NENA with long *ō* in Central Neo-Aramaic in salient stressed syllables but not in non-salient syllables. In other NENA dialects of the Bohtan cluster, such as Hertevin, a long *a* is realized as a back [a:], i. e. with a lesser degree of convergence with *ō*. In Armenian dialects an *a* shifts to *ɔ* in stressed syllables in the area of Cilicia and even to the high vowel *u* in Svedia (Martirosyan, chapter 2.2, §6.1). In Kurdish dialects of the area of Mardin a low central unrounded [a:] of Standard Kurdish is retracted to [a:] and, in dialects further west, has the quality of [ɔ:] (Haig, chapter 2.3, §3.1.2).

Although the diffusion of the backing and rounding of *ā* is not a general feature of the sound systems of NENA dialects, it can be identified in the paradigm of the possessive suffixes in all dialects, where it has been exploited as a strategy of distinguishing otherwise homophonous suffixes. The 2fs suffix is *-ax*. The historical form of the 2ms suffix is **-āx* and by the normal process of historical phonology the reflex of this in the modern NENA dialects should have been *-ax*, i. e. a homophone of the 2fs suffix. In order to resolve this homophony the long **ā* of the historical 2ms form **-āx* shifts to */o/*, which results in the maintenance of the distinction between the 2ms *-ox* and 2fs *-ax* in the paradigm. This shift of **ā > /o/* in the 2ms suffix *-ox*, which, as remarked, is not a general feature of all NENA dialects, can be identified as a convergence with the phonology of neighbouring languages that is motivated by the needs of the morphological system.⁹ In such cases the areal feature is activated beyond the area of general diffusion to enrich the possibilities of making distinctions between linguistic forms. Another example of this phenomenon can possibly be identified in the limited diffusion of pharyngeals in Kurdish, where its sporadic use in native words in some cases appears

⁹ Cf. the work of Malkiel (1968, 1976) on the morphological motivations for “irregular” sound changes in Romance.

to have a similar function of making clear distinctions in a paradigm, e. g. *haft* ‘seven’, *hašt* ‘eight’ (see above).

The provisional conclusion we draw from the Western Asian findings is that in contact situations marked phonological segments differ in their ease of borrowability, and in the degree to which they are retained in languages that have them originally. It seems that glottalized consonants can quite readily cross language boundaries and be incorporated into the native lexicon of the recipient language. Furthermore, these segments seem to be retained in the donor languages. Inter-dental consonants, on the other hand, are not borrowed, and are frequently lost in languages that originally had them. Vowel shifts such as the fronting of rounded vowels and the backing and rounding of *ā* diffuse in different degrees. Salience of sounds is a factor that is conducive to diffusion. Finally, language-internal factors, such as the need to make paradigmatic distinctions, may also bring about a convergence in specific morphological or lexical items.

3.3. Morphosyntax

3.3.1. The marking of definiteness

A case of areal diffusion of morphology is the spread of the Iranian definite marking nominal suffix *-aka*, native to Central and Southern Kurdish and Gorani, to unrelated languages of the region. This exhibits various degrees of integration in the recipient languages. In the Turkic dialects of the region (Bulut, chapter 3.5, §2.4.4.1, chapter 4.2, §2.4.2.2) it is added, as in Kurdish and Hawrami, directly after the stem, before case and plural suffixes, e. g. *šär-äkä-si-ni* (poem-DEF-POSS.3SG-ACC) ‘this certain/aforementioned poem of his’ (Sonqorî). Compare for example Hawrami, where it is followed by number and gender inflection (Mahmoudveysi and Bailey, chapter 4.5, §4.1.1), e. g. *warg-aka* (wolf-DEF.M) ‘the wolf’, *adä-kē* (mother-DEF.F) ‘the mother’, *palawar-akē* (bird-DEF.PL) ‘the birds.’ In NENA dialects, however, it has the fixed form *-äke*, which corresponds to the oblique inflection of the Kurdish particle (*-akay*) rather than the nominative form (*-aka*), e. g. Jewish Sanandaj *bela* ‘house’, *beläke* ‘the house’; Christian Sanandaj *besa* ‘house’, *besäke* ‘the house’ (Khan, chapter 3.4, §6.1, chapter 4.4, §6.3.1). Moreover in NENA it is attached to the right periphery of the noun after plural suffixes, e. g. Jewish Sulemaniyya *barux-awal-äke* (friend-PL-DEF) ‘the friends’, contrast Kurdish *dost-ak-än* (friend-DEF-PL) ‘the friends’ (Khan 2007: 201–202). The stress patterns in the NENA dialects reflect the fact that this definite particle has the prosodic status of an affix rather than a clitic.

Another language that has developed a novel marker of definiteness is the Iranian language Kumzari (van der Wal Anonby, chapter 4.7), spoken in the extreme south of the region. The marker is a suffix with the form *-ō* (*qiṣr* ‘palace’, *qiṣr-ō* ‘the palace’, van der Wal Anonby, chapter 4.7, §4). The origin of the marker

itself is obscure, but it seems likely that the development of systematic marking of definiteness, otherwise rare in Iranian, was influenced by contact with Arabic, which generally employs definiteness marking. The languages to the north of the region lack definiteness markers (Ossetic, Iranian languages of the Caspian, Laz, NENA dialects in the area of Northern Kurdish). Romeyka, Armenian and Central Neo-Aramaic, however, have inherited morphological markers of definiteness. A ‘definiteness isogloss’ across the region as a whole thus cannot be drawn, but a general tendency is that the languages to the southeast of the region are more likely to mark definiteness than those northward of Lake Van.

3.3.2. Commonalities in pronominal and case morphology

In various Anatolian Arabic dialects, plural pronouns and pronominal inflections of common gender contain a *n* element. In such forms one would expect *m*, which is a feature of the masculine plural pronouns in other Arabic dialects and Old Arabic, e. g. Mardin *hənne* (3CPL independent pronoun), *-hən* (3CPL pronominal suffix), *'əntən* (2CPL independent pronoun), *-kən* (2CPL pronominal suffix). It is likely that these were influenced by the form of the corresponding pronouns in the Neo-Aramaic dialects of the area, which contain *n*, e. g. Turoyo (Mardin area) *hənne* (3CPL independent pronoun), NENA Hertevin (Bohtan area) *'ahniton* (2CPL independent pronoun), *-ehon* (2CPL pronominal suffix) (Khan, chapter 2.5, §6.1).

Some NENA dialects spoken in the region of Northern Kurdish have developed a series of demonstrative pronouns from inherited morphology that resembles very closely the shape of corresponding demonstratives in Kurdish, e. g. NENA Barwar (Khan chapter 3.4, §6.1) *'awwa* ‘this’, *'aw* ‘that’ (anaphoric), *'awaha* ‘that over there’, compared with Kurdish of Amedia (MacKenzie 1961: 82, 174): *awa* ‘this’, *aw* ‘that (anaphoric), *awēhē* ‘that over there’. In some NENA dialects in contact with Turkic the masculine singular far deixis demonstrative pronoun has developed the form *'o* ‘that’ by a process of contraction from **'aw* and, moreover, it has supplanted the feminine singular form, with the result that *'o* is now of common gender. This resembles the form and distribution of its demonstrative counterpart *o* in Turkic, which, as in general in Turkic, is genderless (Bulut, chapter 4.2, §2.3.2).

The configuration of demonstrative systems and their syntax converge in various areas of the region. Several languages in the region of Anatolia and northern Iraq, for example, have a three-way configuration of the demonstrative system including (i) near deixis, (ii) anaphoric and (iii) far deixis. This is found in Northern Kurdish and neighbouring NENA dialects (examples from Kurdish of Amedia and NENA of Barwar are given above). This three-way distinction of demonstratives is found in Eastern Armenian (Dum-Tragut 2009: 129–130). In other areas of the region a two-way demonstrative system predominates consisting of near deixis and far deixis forms.

A case of the convergence of the syntactic patterns of demonstratives is as follows. In the Northern Kurdish dialects demonstrative constructions typically consist of a demonstrative pronoun before a noun and a postposed deictic clitic, e. g. *ew-defter=e* ‘that book’, *ev-defter=e* ‘this book’ (Haig, chapter 2.3, §2.2.3). This pattern has been replicated by Central Neo-Aramaic, in which demonstrative constructions consist of a preposed definite article (originally an anaphoric pronoun) and a postposed deictic suffix, e. g. *'u-malk-ano* ‘this king’ (Khan, chapter 2.5, §6.1).

The patterns of genitive constructions converge across various languages (Khan, chapter 2.1, §2.2.1). In genitive constructions in noun phrases in the NENA dialects, for example, the genitive particle is typically suffixed to the head noun, whereas historically it was attached to the front of the dependent noun, e. g.

- (1) Christian Urmi
bet-ət malca
 house-GEN king
 ‘the house of the king’

This is likely to be a convergence with the pattern of the Northern Kurdish attribute marking *ezafe*, which is suffixed to the head noun, e. g.

- (2) a. *xani-yê mirov-î*
 house-EZ.M man-OBL.M
 ‘the man’s house’
 b. *xani-yê wî mirov-î*
 house-EZ.M that.OBL.M man-OBL.M
 ‘that man’s house’

In NENA there is even a replication of the oblique case morphology of Kurdish. As seen in (2) dependent nouns, pronouns, and demonstratives in Northern Kurdish have oblique case morphology. In Neo-Aramaic the oblique morphology of the demonstratives has been replicated by the development of an innovative oblique pronoun through the bonding of the genitive particle with the stem of the demonstrative resulting in an unitary morphological form, e. g. ms. *do* < **d-* (GEN) + *o* (DEM). This is used together with a genitive particle suffixed to the head noun, thus replicating the Kurdish pattern head noun-*ezafe* + oblique pronoun:¹⁰

- (3) C. Urmi
bet-ət do malca
 house-GEN that.OBL king
 ‘the house of that king’

¹⁰ For a detailed treatment of this process see Gutman (2018).

3.3.3. Clitic pronouns indexing verbal objects

One of the most salient features of Iranian languages of western Iran is the high frequency of clitic pronouns, used in several different functions including adnominal possession, prepositional complement, and various kinds of objects (see e. g. Haig, chapter 3.2 for Central Kurdish, Mahmoudveysi and Bailey, chapter 4.5, for Hawrami, Paul, chapter 4.6, §4.9.1 for Persian). When indexing objects, they commonly cliticize to the finite verb. A number of non-Iranian languages have converged with this pattern of object marking. Some incipient cases of this can be identified in Turkic varieties in Iran, in which the use of possessive suffixes on nouns has been extended to the expression of benefactive and dative pronominal arguments on predicative constructions, e. g. *lazim-miz* (necessary-POSS.1PL (is)) ‘we need’, *errebât var-i* (relation existent-POSS.3SG), literally ‘his contact (is) existent’, that is: ‘he has contact’ under the influence of Iranian languages (Bulut, chapter 4.2, §2.4.2.3). In the Neo-Aramaic dialects of western Iran there has been an extension of the use of possessive suffixes to express pronominal direct objects in various parts of verbal paradigms (Khan, chapter 4.4, §7.1, §7.2). In some Armenian dialects in contact with West Iranian languages, possessive suffixes are likewise used on verbs, e. g. Urmî dialect: *me ci p̄n̄enk* ‘-t (one horse let’s. catch-2SG) ‘let’s catch a horse for you’ (Martirosyan, chapter 2.2, §6.6). Similarly, in Iraq Turkic and Iran Turkic, the actual Iranian pronominal clitics (matter borrowing) are used to index benefactives, and direct objects on the Turkish verb, as in *yäyipti=šan* (eat.PF.3SG=3PL) ‘he ate **them**’ (Bulut, chapter 3.5, §2.4.4.2).

3.4. Constituent order

3.4.1. Position of copular elements

For clauses in the present indicative with non-verbal predicates, all languages of the region have an overt, clause-final, obligatory copula element (often clitic in nature), either as one variant, or as the sole option. The general pattern for such clauses is schematically provided in (4a) and (4b):

- (4)
 - a. house large=is ‘the house is large’
 - b. this my.son=is ‘this is my son’

For the Iranian languages, an obligatory clause-final copular is an inherited feature which characterizes most (perhaps all) of contemporary West Iranian, including Iranian languages outside of the region under consideration. Example (5) from Gilaki (Caspian region, see Stilo, chapter 5) is illustrative of contemporary Iranian languages, and is basically matched by all the Iranian languages of the region (see chapters 2.3, 3.3, 4.3, 4.5, 4.6, 4.7 and 5 for further illustration of Iranian copular constructions).

- (5) Gilaki, Lāhijāni (Stilo, chapter 5, ex. 01)
- i ti pəsér=ə*
this your son=COP.3SG
'Is this your son?'

Examples from other language groups are as follows:

- (6) Armenian (Hamšen dialect, Martirosyan, chapter 2.2, ex. 8)

im ɔnun-s Ašot ä
I.GEN name-my PN(MASC) COP.3SG
'My name is Ašot.'

- (7) Arabic of northern Iraq, dialects east of the Tigris (Procházka, chapter 3.2, ex. 1)

báyt-a-hu
house-3SG.F-COP.3SG.M
'It is her house.'

- (8) Turkic of Iran (Khalaj of Bayâdestân, Bulut, chapter 4.2, ex. 10)

her biri älli metir=dir
every one.POSS.3SG 50 meter=COP.3SG
' [...] each one of them is fifty meters (long)'

- (9) NENA of Iran, Jewish dialect of Sanandaj (Khan, chapter 4.4, ex. 13)

tat-óx tajär=ye? mam-i tajär=ye lá tat-i.
father-your merchant-COP.3MS uncle-my merchant-COP.3MS NEG
father-my
'Is your father a merchant?' 'My uncle is a merchant, not my father.'

- (10) Neo-Mandaic of Khorramshahr, Iran (Khan, chapter 4.4, ex. 18)

man=ye?
who=COP.3S
'Who is he?'

- (11) Laz (Lacroix, chapter 6.2, ex. 45)

uʃkuri tʃk'om-eri t'u
apple eat-PART be.IMPFT.13SG
'The apples were eaten.' (state)

Note that Laz has an inflected copular verb. Positionally, however, it matches the clitic copular elements of the other languages considered here.

- (12) Romeyka (Schreiber, chapter 6.4, ex. 8)

<i>Ato</i>	<i>to</i>	<i>kitap</i>	<i>temo</i>	<i>en.</i>
3SG.N	DEF.ART.N	book	1SG.POSS	be.3SG

‘This book is mine.’

The clause-final copula thus appears to be a good candidate for a syntactic trait that has diffused across the entire region. As mentioned, for the Iranian languages, a clause-final copula going back to the Old Iranian defective verb ‘be’ is widespread across the family, and its presence in all Iranian languages of the region is thus a matter of historical inheritance. For the other languages, the obligatory clause-final copula has emerged through quite varied means. Across Turkic generally, a zero-copula is widespread for the third person present indicative, but an overt copula is available as a pragmatically marked option (e. g. in Standard Turkish *dIr*, going back to a verb *dur-* ‘stand’). In the varieties of Turkic investigated here, this marked option has become regularized, so that most present indicative third person copular clauses have an overt copula. Thus for Turkic, the (near obligatory) copula has arisen through relaxing the pragmatic conditions on the old overt copula, so that its frequency of occurrence increases. As Bulut (chapter 4.2, §2.3.3.1) notes, the frequent “usage of the copula of the 3rd person may point to the influence of Iranian languages, where an overt marking is obligatory.” For Semitic languages, the presence of an overt, clause-final copula has required quite extensive restructuring. The morpheme generally identified as the copula in these languages is mostly of pronominal origin, e. g. Anatolian varieties of Arabic (Procházka, chapter 2.4, §2.4.1) and Neo-Aramaic (Khan, chapters 3.4, and 4.4 for NENA, and 2.4 for Central Neo-Aramaic and NENA). Among other things, Khan discusses different degrees of grammaticalization of the innovated copula in NENA. In dialects such as Christian Barwar of northern Iraq, the default position of the copula is clause final, but it may occur on a clause-internal constituent that is in “narrow focus” (Khan, chapter 3.4, §8.1). In other dialects, however, such as Jewish Sanandaj of Iran, the copula element is positionally fixed in clause-final position, and is thus impervious to pragmatic considerations (Khan, chapter 4.4, §8.1.1). The fixed clause-final position matches that of the main contact languages (different varieties of Kurdish), and it is noteworthy that the NENA dialects with fixed clause-final copulas are generally those that have undergone the heaviest contact influence. For Romeyka (Schreiber, chapter 6.4), an overt obligatory copula is part of the Indo-European heritage, but its clause-final position can be presumably attributed to contact influence from Turkish.

In sum, the obligatory clause-final copula in the present indicative is a common template across the region. The different languages have converged on this pattern via varying processes, ranging from major restructuring (re-analysis of pronouns) in the case of Semitic, or relaxation of pragmatic conditions on existing structures (Turkic). Although the processes themselves differ, they nevertheless converge on a common outcome. In this case, the target model has been provided by the

Iranian languages of the region. We may surmise that copular constructions are particularly prone to contact influence due to their omnipresence in everyday communication (e. g. in utterances such as ‘what’s this?’, or ‘that’s mine’, ‘he’s at home’ etc.), and are presumably among the earliest utterance types to be acquired in multi-lingual settings.

3.4.2. The relative ordering of direct object and verb

The relative order of direct object and verb remains one of the most intensely researched features in language typology, and is of correspondingly high interest in areal linguistics. The languages of the region exhibit historically both options: Semitic is historically VO, while Turkic and Iranian are OV across all historically attested stages. For the remaining languages, the issue is more controversial: Greek and Armenian might be considered VO for the earliest records, while old Kartvelian has been characterized as “free SOV/SVO” (Testelec 1998: 236).

In general, the historically attested (or reconstructed) order remains stable in the modern languages of the region. Thus change is the exception, rather than the rule. Contemporary Armenian (Martirosyan, chapter 2.2) and Romeyka (Schreiber, chapter 6.4) are difficult to classify, evidently permitting both orders, and thus seem to continue the somewhat contentious nature of word order in their historical ancestors; clarifying the issues here would raise complications beyond the scope of this introduction. The Kartvelian language Laz (Lacroix, chapter 6.2) has shifted from the apparently free word order of early Kartvelian noted above to a consistent OV order, presumably under Turkic influence. But the most radical examples of change documented in this volume are found in NENA dialects, in particular the trans-Zab Jewish dialects of northern Iraq and northwestern and western Iran. Here the inherited Semitic VO order has completely switched to OV. An example from the Jewish Sanandaj dialect of NENA (Khan, chapter 4.4, ex. 44) is the following:

- (13) *hämər-*Ø *ke* *'ay-brona* ***həl-day*** *brata* *gb-e.*
say.IRR-D.3MS COMP that-boy OBJ-DEM.OBL girl love.PRS-D.3MS
‘[in order that] he say that the boy loves **the girl**’

In Neo-Mandaic, both VO and OV are possible, with the choice apparently dependent on the definiteness and specificity of the object. The contrast is shown in (14), with an indefinite specific object pre-verbally, and (15) with a definite object post-verbally (both from Neo-Mandaic of Khorramshahr):

- (14) *jisr-i* *tum* *əbad-yon*
bridge-INDF again do.PST-3PL
‘They built another bridge.’ (Khan, chapter 4.4, ex. 28, citing Häberl 2009)

- (15) *čāre abd-etton ... qam tā jisər*
 remedy do.IPFV-2PL to that bridge
 ‘You will fix that bridge.’ (Khan, chapter 4.4, ex. 31, citing Häberl 2009)

A contact-induced change from VO to OV is thus clearly documented in NENA, and partially for Neo-Mandaic. A change in the opposite direction, however, is not found in the material covered in this volume. Even Kumzari, an (arguably) Iranian language heavily influenced by neighbouring varieties of Arabic, has largely retained the Iranian OV word order. However, OV order only obtains with lexical objects, while pronominal objects are generally post-verbal, evidently replicating the syntax of Southern Arabian languages. Thus we find that pronominal syntax is apparently more susceptible to contact-induced change than the syntax of lexical noun phrases. Typical examples illustrating pre-verbal lexical objects are given in (16) and (17):¹¹

- (16) Kumzari (van der Wal Anonby, chapter 4.7, ex. 14)

dar-ō twākš-um na.
 door-the open.IMPF-1s NEG
 ‘I will not open **the door**.’

- (17) Kumzari (van der Wal Anonby, chapter 4.7, ex. 46)

sirx dō-um ba yē.
 gold give.IMPF-1s to 3s
 I will give her **gold**.’

In sum, the relative ordering of direct object and verb is a comparatively stable syntactic feature. The only changes attested in the region involve a shift towards OV, rather than in the other direction. This is a particularly interesting finding in view of the widespread opinion that changes from OV to VO (as in Germanic) are in some sense the more natural and frequent kind of word-order change than the

¹¹ The position of pronominal objects in Kumzari is noteworthy because it violates Greenberg’s Universal number 25, according to which if pronominal objects follow the verb, then nominal objects likewise do. A possible account might run as follows: the pronominal objects of Kumzari basically appear in the position of the clitic object pronouns of many West Iranian languages, including Persian, namely immediately following the verb, or the light verb complement if the verb is a complex predicate, cf. examples (4) and (5) in van der Wal Anonby (chapter 4.7). Furthermore, the object pronouns of Kumzari show obvious phonological similarities to the clitic pronouns of other West Iranian languages. What appears to have happened is that the erstwhile object clitic pronouns have developed into free pronouns, but have retained their position to the right of their former hosts. Such a development (clitic pronoun > free pronoun) is obviously unusual, but given that Greenberg’s Universal 25 is a fairly robust generalization, violations presumably require exceptional circumstances such as very heavy contact influence.

opposite direction (see Gell-Mann and Ruhlen 2011 for a recent statement to this effect).

3.4.3. OVG (direct object – verb – goal) word order in the region

While word-order typology has focussed on the position of direct objects relative to the verb, the position of other verbal arguments relative to the verb has received much less attention (see Hawkins 2008 for discussion). The implicit assumption is that generally, all objects will occur on the same side of the verb, thus in VO languages we can expect indirect objects etc. to be post-verbal, and in OV languages we expect them to be pre-verbal. Languages that violate this expectation are quite rare, and interestingly, generally appear to be OV languages (Hawkins 2008). A relatively well-known example is Mande (generally classified as Niger-Congo, West Africa), where adpositional arguments follow, while direct objects precede the verb (Nikitina 2011).

In the Western Asian context, a similar phenomenon occurs in the majority of OV languages: constituents expressing goals of verbs of movement, or of caused movement ('put', 'place' etc.) overwhelmingly follow the verb, yielding a characteristic OVG (G=Goal) order across much of the region (Haig 2014b, 2017, to appear). The generalization that can be drawn is the following:

- (18) Phrases expressing goals of verbs of motion and caused motion are post-verbal, irrespective of the position of the direct object in the language concerned.

For the VO languages of the region, e. g. the Arabic dialects west of the Tigris (Procházka, chapter 2.4), this yields an unremarkable and harmonic word order, in which all manner of verbal complements follow the verb. The following example from Mosul Arabic illustrates the post-verbal position of a goal ('to.house'), and of an indirect object ('some apples'):

- (19) *qabəl-mā y-ǵōh ſa-l-bēt ²štagā təffāh*
before-PRTCL 3SG.M-go.IPFV to-DEF-house buy.PFV.3SG.M apples.COLL
'Before he went **home** he bought **some apples**.' (Procházka, chapter 2.4, ex. 29)

In OV languages, however, (18) leads to a disharmonic order, with direct objects preceding, but goals following the verb. This is illustrated in (20), from the Jewish Sanandaj dialect of NENA. The direct object 'one cow' of 'buy' precedes the verb, while the goal of 'bring' follows the verb:

- (20) [...] *xa-dána* *tórtá* *šäqəl-Ø-wa-la.*¹
 [...] one-single cow buy.PRS.-D.3MS-PST-L.3FS
k-m-e-wa-l-ó *ga-béla* *nòš-ef.*¹
 IND-bring-D.3MS-PST-L.3FS-PVP in-house self-POSS.3MS
 '(each family went and) bought a cow. They would bring it to their
 home.' (Khan, chapter 4.4, taken from ex. 58)

The OVG word order illustrated in (20) is found in all the varieties of Neo-Aramaic that have shifted to OV, and in all the Iranian languages of the region. Example (21) shows post-verbal position of a recipient, coupled with preverbal order of the direct object, in Kumzari:

- (21) Kumzari (van der Wal Anonby, chapter 4.7, ex. 3)
mi̥t̥ dō'a ba diry̥t̥in-an.
 fish give.3S.IMPF to fisher-PL
 'He gives **fish to the fishers**.'

Turkic varieties of the region likewise show a strong tendency to place goals after the verb, thus disturbing the otherwise consistently verb-final syntax (see Bulut, chapter 3.5, §2.4.5.1 for Turkic in northern Iraq, and chapter 4.2, §7.2.4 for Turkic in Iran). The data from Armenian do not permit a firm conclusion on the order of goals. Laz and Ossetic, on the other hand, both seem to be outside of the region where (18) holds. Example (22) from Laz shows that the goal of the verb of caused motion is preverbal, while (23) shows preverbal goals of motion verbs in Ossetic:

- (22) Laz (Lacroix, chapter 6.2, ex. 32)
Xalili-k u/kui dʒebi-s dol-i-by-am-s
 Halil-ERG apple pocket-DAT PV-MID-tip-TH-13SG
 'Halil tips the apples **in his pocket**.'

(23) Ossetic (Erschler, chapter 6.3, ex. 27b)
derxevš-mə səw-ag išči gošt-mə akumulatər
 Darghavs-ALL go-NMZ someone Gosht-ALL accumulator
a-laš-zen
 PRV-carry-FUT.3SG
 'Someone who is going **to Darghavs** will take the accumulator **to Gosht**'

The areal epicentre for OVG word order appears to be northern Iraq and neighbouring regions of western Iran. In Behdinī Kurdish of northern Iraq, post-verbal positioning of goals and recipients is a grammatical rule (Haig, to appear), while in other languages, for example colloquial Persian, or Iran Turkic, it is a statistically preferred option (see Frommer 1981 and Haig 2017 on spoken Persian). Languages also differ in the type of argument that is treated as ‘goal’. Thus in Behdinī dialects of Northern Kurdish goals, recipients, addressees of ‘tell’ (*gotin*), and final states of change-of-state predicates (‘become, turn into’ etc.) are placed

after the verb (Haig, to appear), but dialects of Northern Kurdish further north put addressees of ‘tell’ in front of the verb (cf. Haig, chapter 2.3, Fig. 5). Stilo (2010) notes for a dialect of Vafsi (Tati, see also Stilo, this volume, chapter 5) that recipients of ‘give’ are overwhelmingly post-verbal, while addressees of ‘say, tell’ are overwhelmingly pre-verbal.

Although we lack detailed syntactic analysis for many of the languages in this volume, it is nevertheless possible to formulate some provisional conclusions with regard to post-verbal constituents. Firstly, they are predominant in the region of northern Iraq, western Iran, and southeastern Anatolia. The extent and frequency of post-verbal constituents fades out as one progresses northward and eastward. In Ossetic and Laz, the phenomenon is only marginal. In spoken Persian they are very common, but not grammatically obligatory (Haig 2017), and further eastward (e. g. the easternmost dialects of Balochi, or in East Iranian languages, for example the Wakhi texts in Obřtelová 2017), they are scarce. Second, there appears to be a hierarchy of post-possibility, approximately as follows:

- (24) HIERARCHY OF POST-POSIBILITY IN OV LANGUAGES OF WESTERN ASIA
 (THOSE TO THE LEFT ARE MOST FREQUENTLY POST-POSED)
 goals of (caused) motion > recipients of ‘give’ > addressees of ‘tell’ >
 other

More generally, the wide distribution of OVG order (in Iranian, some of Neo-Aramaic, and Turkic) suggests that goal arguments are comparatively susceptible to “synchronization” with a contact language, leading to a common ordering, while the linear placement of direct objects is more resistant to change (see Haig 2014b).

3.5. Other issues in contact-induced syntactic change

Another area where areal distribution is evidently relevant is adpositions. In general, the Semitic languages have all retained their inherited prepositions, regardless of heavy contact influence. The Turkic languages have likewise largely retained their postpositions. Iranian languages of the region, however, exhibit all possible kinds of adpositions, which approximately correspond to a north-south cline. While Southern and Central Kurdish, Kumzari, Hawrami, and Bakhtiari (and Persian) are overwhelmingly prepositional, the dialects of Northern Kurdish exhibit a mix of pre-, post- and circumpositions. The Caspian and most Tatic languages (Stilo, Chapter 5, §5.7), on the other hand, have a “predominance of postpositions”, as does Ossetic (see Stilo 2005, and Haig, chapter 2.3). Laz and Armenian are likewise predominantly postpositional. Adpositional type is generally slow to change, so the evident variability across Iranian is presumably the result of long-standing convergence with languages of different types.

Relative clauses and subordination generally are also domains known to reflect contact influence (see Gandon 2016). Most of the languages of the region have

post-verbal relative clauses, and this is even found in Iran Turkic, where generally patterns of subordination have moved closer to the Iranian type (see Khan, chapters 2.1 and 4.1 for an overview of subordination). Ossetic has the most complex patterns of subordination, and has also pre-verbal participial relative-like constructions (see the nominalized verb ‘go’ in the Ossetic relative clause in (23) above), perhaps reflecting a shift from the original post-verbal Iranian type towards the pre-verbal pattern in the Caucasus. Laz is generally head-final in this respect, again reflecting the overall trend for the languages to the north of the region to adhere more closely to a consistent head-final type.

3.6. Provisional remarks on ease of syntactic borrowing in the Western Asian context

The data we have considered above, though not a comprehensive survey, do allow provisional conclusions regarding the ease with which the linear ordering of clause constituents can adapt to, or align with, that of a contact language (we have not included patterns of clause linkage and subordination in this hierarchy, but it seems evident that they would rank highly, see Haig 2001, Matras 2002 for discussion). These are summed up in the following hierarchy:

- (25) PROVISIONAL HIERARCHY OF EASE OF CONTACT-INDUCED CHANGE IN
 CLAUSE-INTERNAL SYNTAX
- | | |
|---------------------------------|---|
| copular constructions | ↑ |
| order of goal and verb | ↑ |
| order of direct object and verb | ↑ |
| adpositional order | ↑ |
- HIGHEST
 ↓
 LOWEST

Support for the hierarchy in (25) comes from the blanket distribution of the shared copular construction, the almost complete spread of post-verbal goals (excepting Laz and Ossetic), the small number of unambiguous cases of object/verb ordering changes (e.g. the Trans-Zab Jewish dialects of NENA). But even the languages most heavily affected by contact, such as the aforementioned varieties of NENA, have not abandoned their inherited adpositional type (though of course some of the contact languages are also predominantly prepositional). It may in fact turn out that the relative position of object-verb, and adpositional order in (25) will need to be reconsidered; this remains a task for future research.

References

- Arnold, Werner. 2011. Western Neo-Aramaic. In Stefan Weninger, Geoffrey Khan, Michael Streck & Janet Watson (eds.), *The Semitic languages: An international handbook*, 685–696. Berlin & Boston: De Gruyter Mouton.
- Belleli, Sara. 2016. A study on language and folklore in the city of Harsin (Kermānshāh Province, West Iran). Sketch grammar with texts and lexicon. Naples: University of Naples dissertation.
- Bickel, Balthasar, Johanna Nichols, Taras Zakharko, Alena Witzlack-Makarevich, Kristine Hildebrandt, Michael Rießler, Lennart Bierkandt, Fernando Zúñiga & John B. Lowe. 2017. *The AUTOTYP typological databases*. Version 0.1.0 <https://github.com/autotyp/autotyp-data/tree/0.1.0> (accessed 15 December 2017).
- Blevins, Juliette. 2017. Areal sound patterns: From perceptual magnets to stone soup. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 88–121. Cambridge: Cambridge University Press.
- Butts, Aaron (ed.). 2015. *Semitic languages in contact*. Leiden & Boston: Brill.
- Chirikba, Viacheslav A. 2008. The problem of the Caucasian Sprachbund. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*, 25–94. Amsterdam & Philadelphia: Benjamins.
- Dawkins, Richard. 1916. *Modern Greek in Asia Minor. A study of the dialects of Silli, Capadocia and Phárasa, with grammar, texts, translation and glossary*. Cambridge: Cambridge University Press.
- Dryer, Matthew S. 1988. Universal of negative position. In Michael Hammond, Edith Moravcsik & Jessica Wirth (eds.), *Studies in syntactic typology*, 93–124. Amsterdam & Philadelphia: Benjamins.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138.
- Dum-Tragut, Jasmine. 2009. *Armenian*. (London Oriental and African Language Library, 14). Amsterdam & Philadelphia: Benjamins.
- Fattah, Ismaïl Kamandâr. 2000. *Les dialectes kurdes méridionaux: étude linguistique et dialectologique*. (Acta Iranica, 37). Leuven: Peeters.
- Frommer, Paul. 1981. Post-verbal phenomena in colloquial Persian syntax. Los Angeles: University of Southern California dissertation.
- Gandon, Ophélie. 2016. La relativisation dans une perspective areale: l'aire Caucase – Anatolie de l'est – Iran de l'ouest. Paris: Université Sorbonne Nouvelle – Paris 3 dissertation.
- Gell-Mann, Murray & Merritt Ruhlen. 2011. The origin and evolution of word order. *Proceedings of the National Academy of Sciences of the United States of America* 108(42). 17290–17295.
- Gräwunder, Sven. 2017. The Caucasus. In Raymond Hickey (ed.), *The Cambridge handbook in areal linguistics*, 356–395. Cambridge: Cambridge University Press.
- Güldemann, Tom. 2018. Areal linguistics beyond contact, and linguistic areas in Afrasia. In Tom Güldemann (ed.), *The languages and linguistics of Africa*. (The World of Linguistics, 11), 448–545. Berlin & Boston: De Gruyter Mouton.
- Gutman, Ariel. 2018. *Attributive Constructions in North-Eastern Neo-Aramaic*. Studies in Diversity Linguistics 15. Berlin: Language Science Press.
- Haig, Geoffrey. 2001. Linguistic diffusion in present-day east Anatolia: From top to bottom. In Robert Dixon & Alexandra Aikhenvald (eds.), *Areal diffusion and genetic in-*

- heritance: Problems in comparative linguistics*, 195–224. Oxford: Oxford University Press.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A Construction Grammar approach*. Berlin & Boston: De Gruyter Mouton.
- Haig, Geoffrey. 2014a. East Anatolia as a linguistic area? Conceptual and empirical issues. In Lale Behzadi, Patrick Franke, Geoffrey Haig, Christoph Herzog, Birgitt Hoffmann, Lorenz Korn & Susanne Talabardon (eds.), *Bamberger Orientstudien*, Band 1. Bamberg: Bamberg University Press.
- Haig, Geoffrey. 2014b. Verb-goal (VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 407–425. Piscataway, NJ: Gorgias Press.
- Haig, Geoffrey. 2017. Western Asia: East Anatolia as a transition zone. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 396–423. Cambridge: Cambridge University Press.
- Haig, Geoffrey. To appear. Post-predicate constituents in Kurdish. In Yaron Matras, Ergin Öpener & Geoffrey Haig (eds.), *Structural and typological variation in the dialects of Kurdish*. London: Palgrave Macmillan.
- Hammarström, Harald & Mark Donohue. 2014. Some principles on the use of macro-areas in typological comparison. *Language Dynamics and Change* 4(1). 167–187.
- Hasselbach-Andee, Rebecca (ed.). In preparation. *The Wiley Blackwell Companion to Ancient Near Eastern Languages*. Malden, MA: Wiley-Blackwell.
- Hawkins, John. 2008. An asymmetry between VO and OV languages: The ordering of obliques. In Greville Corbett & Michael Noonan (eds.), *Case and Grammatical Relations: Studies in honor of Bernard Comrie*, 167–190. Amsterdam & Philadelphia: Benjamins.
- Herin, Bruno. 2012. The Domari language of Aleppo, Syria. *Linguistic discovery* 10(2). 1–52. DOI: 10.1349/PS1.1537-0852.A.412, <http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/l/xmlpage/l/article/412> (accessed 17 March 2018).
- Jastrow, Otto. 2011. Turkish and Kurdish influences on the Arabic dialects of Anatolia. *Türk Dilleri Araştırmaları* 21(1). 83–94.
- Johanson, Lars & Éva Csató (eds.). 2006. *The Turkic languages*. London & New York: Routledge Curzon.
- Khan, Geoffrey. 2007. Grammatical borrowing in North-Eastern Neo-Aramaic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 197–214. Berlin & Boston: De Gruyter Mouton.
- Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi*. 4 vols. Leiden & Boston: Brill.
- MacKenzie, David Neil. 1961. *Kurdish dialect studies I*. Oxford: Oxford University Press.
- Maddieson, Ian. 2013a. Glottalized consonants. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/7> (accessed 17 March 2018).
- Maddieson, Ian. 2013b. Presence of uncommon consonants. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/19> (accessed 17 March 2018).
- Malkiel, Yakov. 1968. The inflectional paradigm as an occasional determinant of sound change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for Historical Linguistics*, 20–64. Austin: University of Texas Press.

- Malkiel, Yakov. 1976. Multi-conditioned sound change and the impact of morphology on phonology. *Language* 52. 757–778.
- Matras, Yaron. 2002. Kurmanji complementation: Semantic-typological aspects in an areal perspective. *Sprachtypologie und Universalienforschung / Language typology and language universals* 55(1). 49–63.
- Matras, Yaron. 2007. The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–73. Berlin & Boston: De Gruyter Mouton.
- Matras, Yaron. 2012. *A grammar of Domari*. Berlin & Boston: De Gruyter Mouton.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Nikitina, Tatiana. 2011. Categorial reanalysis and the origin of S-O-V-X order in Mande. *Journal of African Languages and Linguistics* 32. 251–273.
- Obrtelová, Jaroslava. 2017. *Narrative structures of Wakhi oral stories*. Uppsala: Studia Iranica Upsaliensia.
- Stilo, Donald L. 2005. Iranian as buffer zone between the universal typologies of Turkic and Semitic. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 35–63. London & New York: Routledge Curzon.
- Stilo, Donald L. 2010. Ditransitive constructions in Vafsi: A corpus-based study. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook*, 243–276. Berlin & New York: De Gruyter Mouton.
- Stilo, Donald L. 2012. Intersection zones, overlapping isoglosses, and ‘Fade-out/Fade-in’ phenomena in Central Iran. In Behrad Aghaei & M. R. Ghanoonparvar (eds.), *Iranian languages and culture: Essays in honor of Gernot Ludwig Windfuhr*, 3–33. Costa Mesa: Mazda.
- Testelec, Yakov. 1998. Word order in Kartvelian languages. In Anna Siewierska (ed.), *Constituent order in the languages of Europe*, 235–256. Berlin & Boston: De Gruyter Mouton.
- Weninger, Stefan, Geoffrey Khan, Michael Streck & Janet Watson (eds.). 2011. *The Semitic languages: An international handbook*. Berlin & Boston: De Gruyter Mouton.
- Windfuhr, Gernot (ed.). 2009. *The Iranian languages*. London & New York: Routledge Curzon.

2. Eastern Anatolia and northwestern Iran

2.1. Eastern Anatolia and northwestern Iran: overview

Geoffrey Khan

The languages covered by the chapters in this section include the Indo-European languages Armenian and Kurdish and the Semitic languages Arabic and Neo-Aramaic. The varieties of Turkic spoken in the region are treated in Section 4 of this volume, devoted to western Iran (Bulut, this volume, chapter 4.2). The current chapter describes the language situation that existed at the beginning of the 20th century. Some components of the language situation have undergone major changes since this period. This applies in particular to Armenian and Neo-Aramaic, which are now spoken only by small isolated communities in the Republic of Turkey. At the beginning of the 20th century Armenian was spoken throughout the region and Neo-Aramaic was spoken in numerous communities in southeastern Turkey from the region of Tūr ‘Abdīn (Mardin) eastwards. The language situation of the region had undergone various evolutions also at earlier periods. Armenian and Neo-Aramaic have deeper historical roots, which predate the arrival of Arabic, Kurdish and Turkic languages. All the languages of the region in the modern period exhibit a considerable diversity of dialects. The numerous dialects of Armenian are broadly divided into western and eastern Armenian, with the boundary between the two coinciding roughly with the border between eastern Turkey and Armenia and Iran. The Kurdish dialects belong to the northern variety of the language known as Kurmanji and this can be divided into various subgroups according to geographical area. The Neo-Aramaic dialects fall into the Central Neo-Aramaic subgroup spoken west of the Tigris river and the North-Eastern Neo-Aramaic (NENA) subgroup spoken east of the Tigris. The Arabic dialects are divided into the eastern, sedentary-type dialects (often called *qəltu* dialects, from the form *qəltu* ‘I spoke’, which is a distinctive feature of this group), and the western Bedouin type of dialects.

Numerous contact-induced convergences can be identified in the languages described here. A selection of these are noted in this introductory chapter. The languages are attested in a wide range of dialects, classifiable into a variety of geographical subgroups, and so shared areal features generally relate to specific areas of the wider region.

1. Phonology

1.1. Loss of Semitic interdental fricatives

Some dialects of Arabic and Neo-Aramaic preserve interdental fricative consonants, which are a characteristic feature of Semitic. In many dialects of these languages in the region, however, the interdentals have been eliminated to bring their phoneme inventories more in line with the non-Semitic contact languages, which lack interdentals. In several of the *qəltu* Arabic dialects, for example, the interdental fricatives (*d*, *t*, *ð*) have shifted to stops (*d*, *t*, *d*), sibilants (*z*, *s*, *z*), or labio-dental fricatives (*v*, *f*, *v̪*), respectively (Procházka, chapter 2.4, §1.1). In the Central Neo-Aramaic dialect of Mlahso the interdentals *t* and *d* shift to sibilants (*s*, *z*), whereas in the North-Eastern Neo-Aramaic (NENA) dialects of the region the most frequent process is for the interdentals to shift to stops (*t*, *d*). In some NENA dialects the unvoiced fricative *t* is reduced to zero. In the Jewish dialects of north-western Iran the interdentals shift to the lateral *l*.

1.2. Unaspirated stops

Many Armenian dialects of the region have in their phonemic inventory unaspirated unvoiced stops contrasting with aspirated unvoiced stops, which are found also in Classical Armenian and are possibly an inheritance from earlier Indo-European (Martirosyan, chapter 2.2, §4.1). Unaspirated stops are also found in the Kurdish and Neo-Aramaic dialects of the region. Since unaspirated stops are not a feature of Kurdish and Neo-Aramaic dialects further south outside of the Armenian dialect area, it is highly likely that they entered Kurdish and Neo-Aramaic through contact with Armenian.

1.3. Pharyngeals and pharyngealized consonants

Pharyngeal fricatives (*ħ* and *ʕ*) and pharyngealized consonants are a feature of Semitic and occur in the Arabic and Neo-Aramaic dialects of the region.

The Central Neo-Aramaic subgroup, spoken in the region of Mardin (Tûr ‘Abdîn), preserves the pharyngeal fricatives *ħ* and *ʕ* most conservatively, probably due to extensive contact with Arabic, whereas in most NENA dialects they have been lost. A small cluster of NENA dialects in the region Bohtan on the western periphery of NENA adjacent to the Central Neo-Aramaic dialect area have the pharyngeal *ħ* in their sound inventory (Khan, chapter 2.5, §5.3). Pharyngeal fricatives occur in some Kurdish dialects in the area of Mardin. It is possible that they have entered Kurdish through Arabic or Central Neo-Aramaic.

Pharyngealized consonants occur throughout Arabic and Neo-Aramaic dialects in the region (Khan, chapter 2.5, §5.4). They are found also in Kurdish in

the Mardin area, where as just indicated, also pharyngeal consonants have developed, it seems in association with the pharyngealization of the consonants (Haig, chapter 2.3, §2.1.2). In the NENA dialects in the eastern half of the region, the original pharyngealization of the consonantal segments has developed into a suprasegmental feature that takes a whole word as its domain. In the Christian dialects where this suprasegmental feature develops, an original pharyngealized stop *t* [t] becomes reanalyzed as an unaspirated stop *t̪*, which is facilitated by the existence of unaspirated unvoiced stops in their phoneme inventories.

1.4. *h > x*

The shift of the laryngeal *h > x* is a characteristic of the Armenian dialects of the Van and Urmi areas. In the variety of the Christian Urmi NENA dialect spoken in Armenia a historical *h* shifts to the velar fricative /x/ under the influence of a similar shift in the Armenian dialects of the region, e. g. *xada* ‘thus’ (<*hada*>).

1.5. Palatalization

Palatalization of dorsals is found throughout the region. In NENA the process is particularly common in the southeast of the region and is most advanced in the Christian Urmi dialect, in which the stops **k* and **g* shift without conditioning to affricates č [tʃ] and j [dʒ]. Parallels to such advanced palatalization can be found in the Azeri Turkic and Kurdish dialects of the Urmi region. In the NENA Christian Urmi dialect there is complete palatalization of **k* in some words resulting in *y*, e. g. **duktə* > **duytə* ‘place.’ A parallel to this can be found in the Armenian dialects of the Urmi area, in which unaspirated *k* and *g* become *y* and aspirated *k* becomes *hy* before syllable-initial consonants (Martirosyan, chapter 2.2, §6.6). Kurdish dialects in the Şemzînan area exhibit a lesser degree of palatalization of the stops *g* and *k*, which are realized with a palatal offglide preceding front or central unrounded vowels, e. g. Standard Kurdish [kîra:s], Şemzînan [k'îra:s] ‘robe’.

1.6. Rhotics

In the Kurdish dialects of the region two types of rhotic are found, viz. trills [r] or flaps [ɾ]. These can contrast when not in word-initial position, e. g. [pir] ‘much’ vs. [pir] ‘bridge’ (Haig, chapter 2.3, §2.2). In NENA dialects of the region rhotics are likewise either trills or flaps, which can contrast. The trills, which are pronounced with greater muscular tension, are typically also pharyngealized. In some of the NENA dialects of the Tyare area, in addition to the non-pharyngealized flap and the pharyngealized trill a non-pharyngealized retroflex *ɻ* has developed, resulting in a three-way rhotic opposition, e. g. Ashitha: *paja* ‘layer of dust’, *para* ‘ornamental coin’, *para* ‘lamb’.

1.7. $a > o$

The backing of the vowel *a* is a feature that is shared by many of the languages in the western area of the region. In Armenian dialects an *a* shifts to *ɔ* in stressed syllables in the area of Cilicia and even to the high vowel *u* in Svedia (Martirosyan, chapter 2.2, §6.1). In Kurdish dialects of the area of Mardin a low central unrounded [a:] of Standard Kurdish is retracted to [ɑ:] and, in dialects further west, has the quality of [ɔ:] (Haig, chapter 2.3, §3.1.2). In some NENA dialects of the Bohtan area long *a* in non-final open syllables shifts to *o*, e. g. Ruma *xmóra* ‘donkey’ (< **hmārā*). In other NENA dialects of the Bohtan cluster, such as Hertevin, a long *a* is realized as a back [ɑ:]. In Central Neo-Aramaic spoken further west in the area of Ṭūr ‘Abdīn all historical long *a* vowels have the quality of *o*, e. g. Turoyo: *hmoro* ‘ass’ (< **hmārā*).

1.8. Vowel fronting and raising

In the Kurdish dialects in southeastern Anatolia (Şemzînan) a back high vowel [u:] fronts to [y:] (transcribed here as *ü*), e. g. Standard Kurdish *stur*, Şemzînan *stür* ‘thick’. By a pull-chain effect [o:] in such Kurdish dialects is raised to [u:] to fill the gap left by the fronting of original [u:], Standard Kurdish *bo*, Şemzînan *bu* ‘for’. Similar fronting of [u:] is found in the Kurmanji dialects of the Urmi region ([u:] > [y:] > [i:]), in the Kurmanji spoken in northern Iraq (Badini dialect, Haig, this volume, chapter 3.3, §4.1.2) and in Mukri dialects in northwestern Iran ([u:] > [y:]). These same processes are found in the NENA dialects of the southeastern area of the region, e. g. ⁺Mawana: *xabiša* < **habbūšā* ‘apple’, *ruqe* < **rōqē* ‘spittle’. In the Christian Urmi dialect long *ü* is not fully fronted but shifts to a diphthong with a palatal offglide: *xabuyša* ‘apple’ (Khan, chapter 2.5, §5.5.2).

1.9. $w > x$

In many Armenian dialects of eastern areas of the region the *w* of the diphthong *aw* shifts to *x* by a process of fortition before a dental stop or affricate, e. g. *eawt* ‘*n*’ ‘seven’ > Mełri óxtə (Martirosyan, chapter 2.2, §5.1). This process in Armenian may have been a factor in the fortition of the offglide of long *u* to *x* in some NENA dialects of the area, e. g. Christian Salamas *xibuxša* < **habbūša* ‘apple’ (Khan, chapter 2.5, §5.5.2).

2. Morphology

2.1. Pronouns

A salient feature of the Anatolian Arabic *qəltu* dialects is occurrence of *n* in several plural pronouns and pronominal inflections of common gender. In such forms one would expect *m*, which is a feature of the masc. plural pronouns in other Arabic dialects and Old Arabic, e. g. Mardin *hənne* (3CPL independent pronoun), *-hən* (3CPL pronominal suffix), *'əntən* (2CPL independent pronoun), *-kən* (2CPL pronominal suffix). It is likely that these were influenced by the form of the corresponding pronouns in the Neo-Aramaic dialects of the area, which contain *n*, e. g. Turoyo (Mardin area) *hənne* (3CPL independent pronoun), NENA Hertevin (Bohtan area) *'ahniton* (2CPL independent pronoun), *-ehon* (2CPL pronominal suffix) (Khan, chapter 2.5, §6.1).

In the Kurdish of the region demonstrative constructions typically consist of a demonstrative pronoun before a noun and a postposed deictic clitic, e. g. *ew-deftər=e* ‘this book’, *ev-deftər=e* ‘that book’. (Haig, chapter 2.3, §2.2.3). The clitic may also be attached directly to a demonstrative pronoun when it is used independently, e. g. *ew-e* ‘this one’. This pattern may have influenced the demonstrative constructions in Central Neo-Aramaic, which consists of a preposed definite article (originally an anaphoric pronoun) and a postposed demonstrative suffix, e. g. *'u-malk-anō* ‘this king’ (Khan, chapter 2.5, §6.1.).

In most of the Kurdish dialects of the region there are near deixis and far deixis demonstratives, which are distinguish by the contrast of /v/ and /w/, viz. *ev* (near deixis), *ew* (far deixis). In the Şemzînan area the [v] and [w] are not phonologically distinct and so there is only one demonstrative, viz. *ew*. Near deixis is distinguished by attaching the deictic clitic, e. g. *ew-deftər=e* ‘this book’, *ew-deftər* ‘that book’, *ew-e* ‘this one’, *ew* ‘that one’. These constructions and the phonetic shape of the demonstratives have close replicas in some NENA dialects of the area, which are innovative and have most likely developed by a process of convergence with Kurdish, e. g. Ashitha *'awwa* (< *'aw-ha*) ‘this one’, *'aw* ‘that one’ (anaphoric). The forms have Aramaic etymologies but have taken on the phonetic shape of the corresponding Kurdish forms. The same applies to the *qəltu* Arabic dialect of Siirt, in which near deixis demonstratives have an innovative form as a result of convergence with the phonetic shape of Kurmanji, e. g. *āva* < **hādā* (Jastrow 1978, 1: 108).

Some NENA and Arabic dialects of the region have three types of deixis, viz. (i) near to speaker, (ii) near to hearer, (iii) far from both speaker and hearer, e. g. NENA Ashitha (i) *'awwa*, (ii) *'aw*, (iii) *'awha*, Arabic Daragözü (Jastrow 1978, 1: 109) (i) *āza*, (ii) *āk*, (iii) *ukkā*. The hearer deixis form (ii) functions also as an anaphoric pronoun. The far deixis form (iii) is formed by adding a deictic suffix *-ha*. In the case of the NENA Ashitha (iii) form, this clitic is added to the (i) form,

which already has a deictic clitic, resulting in the accumulation of two suffixes (*'awaha* < *'aw-ha-ha*). This three-way distinction of demonstratives is found in Eastern Armenian (Dum-Tragut 2009: 129–130) and also in some Kurmanji dialects in northern Iraq, which resemble closely the three-way set of NENA demonstratives cited above, e. g. Kurmanji of Amedia: *awa* ‘this’, *aw* ‘that (anaphoric), *awēhē* ‘that over there’ (MacKenzie 1961: 82, 174). Bedir Khan and Lescot (1968: 224) also mention the additional demonstrative particle *ha / he*, which attaches to a noun modified by a demonstrative, following the postposed clitic mentioned above: *ev jinik=a ha* ‘this woman’.

2.2. Noun phrase

2.2.1. Genitive markers

In some Arabic dialects of the region a head noun is combined with a dependent noun in a genitive relation by means of a relative particle (Procházka, chapter 2.4, §4.2), e. g.

- (1) Haksöy

<i>səʕēb</i>	<i>lē</i>	<i>baqqāl</i>
owner	REL/GEN	shop

‘The owner of the shop’ (literally: the.owner the.one.of the.shop)

Constructions with such relative/genitive particles are innovative developments in Arabic but are an archaic feature of Aramaic and continue to be used in the Neo-Aramaic dialects of the region. In Central Neo-Aramaic dialects the particle is prefixed to the dependent noun, as in earlier Aramaic, e. g.

- (2) Turoyo

<i>i-barθo</i>	<i>d-u-malko</i>
ART.FS-daughter	GEN-ART.MS-king
‘the daughter of the king’	

The synthetic construction with a relative/genitive particle in Arabic dialects such as Haksöy may have developed by convergence with this Aramaic construction. Arabic dialects also have an independent genitive pronominal element, e. g.

- (3) Qarṭmīn (Jastrow 1978, 1: 125)

<i>əlfaras</i>	<i>dili-ye</i>
the.horse	GEN.PRO-COP.3FS
‘The horse is mine’	

This also is likely to have developed through convergence with Neo-Aramaic dialects, e. g.

- (4) Christian Urmi

'o-suysa diyyi=ilə
the-horse GEN.PRO=COP.3MS
'The horse is mine'

Parallel constructions are found also in Kurdish with independent *ezafe* particles, e. g.

- (5) *va defter=na yê min=in*
DEM.PROX notebook=DEM.PL EZ.PL 1SG.OBL-COP.3PL
'These notebooks are mine'

In genitive constructions in noun phrases in the NENA dialects the relative/genitive particle is typically suffixed to the head noun, e. g.

- (6) Christian Urmi

bet-ət malca
house-GEN king
'the house of the king'

This is likely to be a convergence with the pattern of the Kurdish attribute marking *ezafe*, which is suffixed to the head noun, e. g.

- (7) a. *xani-yê mirov-î*
house-EZ man-OBL
'the man's house'
b. *xani-yê wî mirov-î*
house-EZ that.OBL man-OBL
'this man's house'

In Kurdish dependent nouns, pronouns, and demonstratives are in an oblique case form. In Neo-Aramaic the oblique of the demonstratives have been replicated by the development of an innovative oblique pronoun:

- (8) C. Urmi

bet-ət do malca
house-GEN that.OBL king
'the house of that king'

2.2.2. Adjectives

In many of the Arabic dialects of the region a noun and its attributive adjective are combined in a genitive-type construction, which is shown by the fact that a definite head noun does not have the article (Procházka, chapter 2.4, §4.2).

- (9) Mardin
bənt əl-əkwayys-e
 girl DEF-beautiful-F
 ‘the beautiful girl’

This is an innovation in Arabic which is likely to have developed by convergence with Kurdish, in which a noun is regularly linked to an adjective by the relative/genitive *ezafe* particle, e. g.

- (10) *bajar-ek-î mezin*
 town-INDEF-EZ.M big
 ‘a big town’

Some features of the syntax of adjectives in Kurdish appear to have had an impact on NENA constructions. The use of independent *ezafe* particles to add additional dependents to a head noun and to nominalize an adjective have parallels in NENA, in which a demonstrative pronoun stands in the position of the *ezafe*, e. g.

- (11) Kurdish
bra-yê min ê mezin
 brother-EZ.M 1SG.OBL EZ.M big
 ‘my older brother’
- (12) C. Urmi (Khan 2016, 2: 49)
brât-malca 'e +jurta
 daughter-king DEM.FS. big
 ‘The elder daughter of the king’

- Kurdish: *yê spî*
 NENA (C. Urmi): *'o +xvara*
 ‘the white one (masc. sing.)

2.3. Prepositions

Some grammaticalization pathways of nouns to prepositions that are found in Kurdish have been replicated in NENA, e. g.

- Kurdish: *ser* ‘head’ *li ser* ‘on, upon’
 NENA (C. Urmi): *riša* ‘head’ *b-rîš* ‘on, upon’

2.4. Attention-drawing particle d-

Several languages of the region use a particle *dV-*, with various different vowel qualities filling the V position, to intensify an imperative. This is found in Arabic dialects, e. g. Mardin: *də-qūmu* ‘stand up (pl.)!’, Harran-Urfa *d-init-ni* ‘give (m.sg.) me!’, *du-gul-li* ‘tell (m.sg.) me!’, Kurmanji Kurdish, e. g. *de bēje* ‘go on, say!’, and Neo-Aramaic dialects, e. g. Ashitha *de-plut!* ‘go out!’. It is possible that it has its origin in an Arabic demonstrative particle *da-/dī* (Procházka, chapter 2.4, §6). This is likely to be related to the Arabic near deixis demonstrative *hādā*. The particle may also be identified as a prefix in the deictic copula of some NENA dialects of the area, in which it typically expresses near deixis, e. g. Christian Urmi *d-ulə* ‘here he is’ (Khan, chapter 2.5, §6.1).

2.5. Verbal constructions

2.5.1. Preverbal particles

All languages of the region express TAM distinctions by affixes attached to verbal stems. There are conspicuous similarities in the patterning of some of these particles across the languages. This is particularly the case with regard to the prefixes expressing the indicative present, which are found in Arabic *qəltu* dialects, Neo-Aramaic, Kurdish and Armenian, e. g.

Arabic (Azəx):	<i>kū-nəktəb</i>	‘we are writing’
Neo-Aramaic (Turoyo)	<i>ko-nošeq</i>	‘he kisses’
NENA (Christian Urmi)	<i>ci-šatə</i>	‘he drinks’
Armenian (Muš)	<i>kə-sirim</i>	‘I love’
Kurmanji	<i>di-xum</i>	‘I eat’

It appears that most, if not all, of these preverbal particles are in origin deictic particles that draw attention to an activity in progress in the present.¹ It can be established that several of them have developed from a deictic particle with a copula enclitic, e. g. Arabic *k-u* < *k-uwe* ‘it is that’, C. Urmi *c-i-* < *c + i*. This probably also applies to the Turoyo prefix *k-o* < *k-yo* or *kale-yo*. One of the etymologies proposed for the Armenian *kə-* is a deictic particle ‘behold’ (Makaev 1977). It is notable that many of the particles have a similar phonetic shape across the various languages, although each is likely to be an internal development. In some lan-

¹ Matras (2010: 75) considers the parallels among the languages discussed here to be the result of a common grammaticalization pathway, suggesting that the prefixes “usually derive from a preposition indicating location or similarity”. This appears unlikely for the languages of Anatolia, but note that the etymology of the Kurdish indicative prefix *di-*, or the functionally equivalent forms *da-*, *a-* in other varieties of Kurdish, has not yet been satisfactorily resolved.

guages, such as Arabic, the progressive is still the basic function of the form. In most languages the aspectual range has been extended to habitual as well as progressive (e. g. Turoyo, Kurmanji and many western Armenian dialects). In these languages a progressive is sometimes specifically marked by additional particles, e. g. Turoyo *kal-ko-nošəq* ‘he is kissing’, or in some cases by repeating the particle, e. g. Armenian *Zeyt'un g> g-arnum* ‘I am taking’. In Southeastern Kurmanji (Bahdini) a progressive is expressed by combining the present with an *ezafe* particle, which functions as a copula in these dialects, e. g. *yē di-hêt* ‘he is coming’.

In Armenian, Neo-Aramaic and Arabic the unrealis subjunctive is expressed by the verb without a preverbal particle, whereas in Kurmanji the subjunctive is expressed by a dedicated particle.

In the eastern Armenian dialects in the east of the region the indicative present is expressed by other types of constructions, mostly consisting of an infinitive (sometimes in the locative case), participle or verbal noun and the copula (Martirosyan, chapter 2.2, §4.2). In NENA dialects east of Bohtan, such constructions based on the infinitive are used to express the progressive, constructions with the preverbal particle being largely restricted to the habitual. In the far east of the region, in northwestern Iran, the usage of constructions based on the infinitive are extended to include the habitual, in some cases as an alternative to the habitual expressed by a construction with a preverbal particle, e. g. Christian Urmi, and in some cases as the regular form for both progressive and habitual, e. g. Jewish Salamas. The distribution of the construction in the Jewish Salamas NENA dialect, therefore, exhibits the greatest convergence with eastern Armenian dialects. The indicative construction based on the copula and infinitive is found also in north-west Iranian languages belonging to the Caspian and Tatic groups (see Stilo, this volume, chapter 5) in adjacent areas in the Caspian region in northern Iran and southern Azerbaijan, where eastern Armenian dialects were also spoken. In some cases, this is the general construction for the indicative, e. g. North Talyshi (Tatic), in other cases it has specifically a progressive sense, e. g. Harzani, Kalāsuri (Caspian) (Stilo, this volume, chapter 5, §6.5.1.3). Preverbal particles are used to form the indicative in north-west Iranian languages spoken further south (Stilo, this volume, chapter 5, §6.5.1.1). Standard Turkish spoken in the Republic of Turkey expresses the progressive by locative constructions based on the infinitive, but the various non-standard Turkic vernaculars of the region do not exhibit this feature (Bulut, this volume, chapter 4.2, §2.3.3.2).

In some Armenian dialects other constructions are used to express the progressive, such as possessive constructions with the verb ‘to have’ (Martirosyan, chapter 2.2, §4.2.), parallels for which can be found in north-west Iranian languages (Stilo, this volume, chapter 5, §6.5.8).

2.5.2. Future

The future tense is expressed by a preverbal particle in most languages of the region. In the Arabic *qəltu* dialects this is *ta-*. In Turoyo (Central Neo-Aramaic) the future is expressed by a preverbal particle with various allomorphs (*g*, *k*, *gd*, *kt*), all of which derive historically from a particle with the form **gəd* (Jastrow 1985: 148–149). This can also express habitual aspect. In Mlaħso (Central Neo-Aramaic) the future particle has the form *d-*, which is likely to be related (Jastrow 1994: 52). It is likely that the alveolar element *t/d* in these forms is a subordinating particle. In NENA dialects of the region the future preverbal particle has the form *bəd-*, which is often contracted to *b-* or, before vowels, *p̪t-* *t̪-*. In Kurmanji Kurdish the future is expressed by the preverbal particle (*d*)*ē* or *wē* followed by the verb with a subjunctive prefix *bi-* (though in the southeastern dialects of Kurmanji, the verb lacks this prefix in the future tenses). The Arabic *ta-* is derived from the purpose particle **hattā* ‘so that, in order that’. The NENA future particle appears to be derived from the volitive expression **bā’ē d-* ‘wants to’. The Kurmanji future marker *dē/wē* is likely to be a fossilized third singular form of the verb *vīyan* / -*vē* ‘be necessary, desirable’, still used in many dialects to express ‘want’, cf. Sorani *awē*, southeastern varieties of Kurmanji *divē(t)*. A parallel to the Turoyo future can be identified in the expression *ga d-* (< **ka d-*) to express the imminent future (‘about to’), which is used in NENA dialects in northern Iraq (Khan 2008: 583–584) and some of the dialects of Anatolia.

In Standard Kurmanji the future particle *dē/wē* etc. can be combined with a negated subjunctive form of the verb to express negated future. However, in the varieties of Kurmanji spoken in Şemzînan and northern Iraq, this is not possible. Instead, a negated form of the indicative is used to express the negated future. The lack of a future particle in negative future constructions is the norm in the *qəltu* Arabic dialects (Jastrow 1978, 1: 313) and the NENA dialects.

In some of the languages of the area future verb forms may also be used to express the habitual aspect. This is the case in Turoyo, NENA and the Kurmanji dialects in the southeast of the region (known as Badini).

2.5.3. Perfect

The perfect is formed in Arabic *qəltu* dialects and in Turoyo Neo-Aramaic by combining a deictic/presentative particle with the past perfective form (Procházka, chapter 2.4, §3.3; Khan, chapter 2.5, §7). The particle is typically the same, or similar to, the one used with present verbs to express the present indicative or present progressive and is a reduced form of a 3s deictic copula, e. g. Arabic Āzəx *kū-ṣərtu* ‘I have become’ (Jastrow 1978, 1: 307), Turoyo *ko-qayəm* ‘he has arisen’.

Armenian and the NENA dialects east of Bohtan express the perfect by a participle and copula. In Armenian this is a perfect participle, e. g. Eastern Armenian

gnac’el em (go.PRF.PTCP COP.1S) ‘I have gone’ (Dum-Tragut 2009: 222), in NENA it is expressed by a participle that functions as both a resultative and perfect participle, e. g. Christian Urmi *xišə-vən* (go.RES.PTCP COP.1MS) ‘I have gone’. In some Armenian dialects the construction with the perfect participle has come to express the perfect, Kak‘avaberd (eastern Armenian), *xəməl əm* ‘I drank’. In Kak‘avaberd the perfect is now formed by a construction with a resultative participle (Muradyan 1967: 139).

The perfect in Standard Kurmanji appears to be a hybrid between these two strategies. Most of the paradigm of Standard Kurmanji looks like a simple past combined with a present 3s copula, e. g. *hat-im-e* (come.PST-1S-COP.3S) ‘I have come’ (cf. *hatim* ‘I came’) but the 3s form looks more like a perfect participle + copula *hatiy-e* ‘he has come’ (cf. *hat* ‘he came’) (Thackston 2006: 53).

In Southeastern Kurmanji dialects the perfect is formed by a combination of the *ezafe* particle with a past participle, e. g. *yē hatî* (EZ.MS come.PST.PTCP) ‘he has come’. In such dialects the *ezafe* functions as a copula, so this is equivalent to the strategy of combining the participle with a copula.

The Mlahso dialect of Central Neo-Aramaic and the NENA dialects of the Bohtan area express the perfect by attaching to a past stem a set of person suffixes that are characteristic of the inflection of the present, e. g. Ruma (Bohtan): *qəm-la* (perfective < *qim* rise.PST-ERG.3FS) ‘she stood up’, *qim-a* (perfect: *qim* rise.PST-DIRECT.3FS) ‘she has stood up’; cf. *qem-a* (present: stand.PRS-DIRECT.3FS) ‘she stands up’.

The western Bedouin subgroup of Arabic dialects express the perfect by the active participle, which is a common strategy across many Arabic dialects (Procházka, chapter 2.4, §3.3).

The languages of the region use the perfect not only to express events from a reference point of the present (resultative or anterior) but also to express the evidential.

2.6. Copula

All the languages of the region express the copula by an enclitic that is attached to the predicate. In the Indo-European languages, Armenian and Kurmanji, the copula is verbal in origin, though phonetically reduced to little more than person markers. In the Semitic languages, Arabic and Neo-Aramaic, the copula is pronominal in origin. In Arabic the enclitic copula is found specifically in the *qəltu* dialects, which share many other areal features with contact languages. In these *qəltu* dialects the copula is clearly pronominal (Procházka, chapter 2.4, §4.1). Also in Central Neo-Aramaic (Turoyo and Mlahso) the enclitic copula is clearly pronominal. In NENA dialects, however, the copula reflects some feature of verbal inflection, which is an innovation that has presumably developed by convergence with the verbal copulas of Kurmanji and Armenian.

In addition to the enclitic copula, *qəltu* Arabic and Neo-Aramaic also have deictic copulas, which draw attention to an entity or a situation. These typically consist of a deictic pronominal element and an enclitic copula, e. g. Arabic (Kəndērīb) *kūwe* (< *k* + *uwe*) (Jastrow 1978, 1: 139), NENA Ashitha *hole* (< *ha-* 'aw=ile) ‘there it is him, there he is’. In Turoyo it consists of an object pronoun rather than an enclitic copula, e. g. *kalé* ‘see him, there he is’.

In the Southeastern Kurmanji dialects (Bahdini) the *ezafe*, originally a relative/attributive particle, has come to be used as a copula. Unlike the enclitic copulas just described, the *ezafe* copula is placed before the predicate, sometimes in an independent form, e. g. *Ergin yê bərsî* ‘Ergin is hungry’. This appears to have developed first in affirmative contingent and locative predicates. Such constructions are replicated in *qəltu* Arabic and NENA dialects by deictic copula constructions, e. g.

- (13) Arabic Qarṭmīn (Jastrow 1978, 1: 140)
əbnu kū qəddām əmm-u
 son-his DEIC.COP.3MS before mother-his
 ‘His son is in front of his mother’

- (14) NENA Ashitha
bron-e hole qam-yəmm-e
 son-his DEIC.COP.3MS before-mother-his
 ‘His son is in front of his mother’

In negative copula constructions, the enclitic copulas are typically attached to the negators in Arabic, Neo-Aramaic and Armenian, e. g.

- (15) Arabic Mḥallami (Procházka, chapter 2.4, §5.3)
ma-na hawn
 NEG-COP.1S here
 ‘I am not here’

- (16) NENA (Ashitha)
lele (< la=ile) laxxa
 NEG-COP.3MS here
 ‘He is not here’

- (17) Eastern Armenian (Dum-Tragut 2009: 215)
Anuš-ě gelec'ik aljik ē
 Anuš.NOM-the beautiful girl.NOM COP.3S
 ‘Anuš is beautiful girl’

- (18) *Anuš-ě gelec'ik aljik č'-ě*
 Anuš.NOM-the beautiful girl.NOM NEG-COP.3S
 ‘Anuš is not a beautiful girl’

3. Syntax

3.1. Alignment of transitive clauses

The Kurdish dialects of the region exhibit an essentially ergative alignment in perfective and perfect transitive clauses, in that the verbal inflection of perfective verbs and the copula of perfect constructions agree with the object of the clause and the agent has oblique case. The Neo-Aramaic dialects of the region exhibit some degree of convergence to the alignment patterns of Kurdish. Subjects of perfective transitive clauses have oblique agreement suffixes in NENA. In many dialects, moreover, a pronominal object is expressed by an absolutive suffix and a definite object argument in the clause is cross-referenced by such an absolute suffix. In some Neo-Aramaic dialects, however, objects are not expressed by absolute suffixes and the alignment is closer to an accusative typology than an ergative one.

The other languages of the region all have accusative alignment.

3.2. Relative clauses

Both dialect groups of Arabic in the region exhibit an innovation in the structure of relative clauses whereby the head noun is put into attributive annexation (“construct state”) with what follows. This is presumably due to convergence with Kurdish *ezafe* constructions, which is used to join both a head to a genitive nominal attribute and a head noun to an attributive relative clause (Procházka, chapter 2.4, §5).

The Southeastern Kurmanji (Bahdini) predicative *ezafe* functions both as copula and as a relative copula, e. g.

- (19) *meriv yê l-mal*
 man EZ.M in-house
 ‘The man is in the house’ ~ ‘The man who is in the house’

The relative function the *ezafe* is replicated in *qəltu* Arabic and NENA by a relative particle with a clitic copula: Arabic (Qarṭmīn) *la-wwe* (REL-COP.3MS), NENA (Ashitha): *t-ile* (REL-COP.3MS), e. g.

- (20) NENA (Ashitha)
 ‘*o-naša t-ile gu-beθa*
 that-man REL-COP.3MS in-house
 ‘The man who is in the house’

References

- Bedir Khan, Emir Djeladet and Roger Lescot. 1968. *Grammaire kurde (dialecte kurmanji)*. Paris: Maisonneuve.
- Dum-Tragut, Jasmine. 2009. *Armenian*. (London Oriental and African Language Library 14). Amsterdam & Philadelphia: Benjamins.
- Jastrow, Otto. 1978. *Die mesopotamisch-arabischen Qəltu-Dialekte*, vol. 1. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1985. *Laut- und Formenlehre des neuaramäischen Dialekts von Mīdin im Tūr ‘Abdīn*. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1994. *Der neuaramäische Dialekt von Mlaḥsō*. (Semitica Viva 14). Wiesbaden: Harrassowitz.
- Khan, Geoffrey. 2008. *The Neo-Aramaic Dialect of Barwar*, 3 vols. Leiden & Boston: Brill.
- Khan, Geoffrey. 2016. *The Neo-Aramaic Dialect of the Assyrian Christians of Urmī*, 4 vols. Leiden & Boston: Brill.
- MacKenzie, David Neil. 1961. *Kurdish Dialect Studies*. London Oriental Series, vols. 9–10. Oxford: Oxford University Press.
- Makaev, E. A. 1977. The origins of the modal particle *ke* in Armenian. In N. Ya Gabaraev (ed.), *Questions of Iranian and General Philology*, 195–200. Tbilisi: Metsniereba publishing house. [In Russian].
- Matras, Yaron. 2010. Contact, convergence, and typology. In Raymond Hickey (ed.), *The handbook of contact linguistics*, 66–85. Maldon & Oxford: Blackwell.
- Muradyan, H. D. 1967. *Kak ‘avaberdi Barbarə*. Yerevan: Academy Press.
- Thackston, Wheeler. M. 2006. *Kurmanji Kurdish: A reference grammar with selected readings*. Unpublished Manuscript, Harvard University. <http://www.fas.harvard.edu/~iranian/Kurmanji/index.html> (accessed 15 January 2015).

2.2. The Armenian dialects

Hrach Martirosyan

1. Introduction

The Armenian language is known to us from the fifth century CE onwards thanks to an unbroken literary tradition, which involves three periods: Classical (5th to 11th centuries), Middle (12th to 16th), and Modern (17th to 21st). Modern Armenian was canonized in the 19th century and consists of two branches: Western (based on western dialects such as that of Constantinople/Polis) and Eastern (based on the dialects spoken on the Ararat plain and surroundings). Armenian is genetically related to Indo-European languages such as Hittite, Sanskrit, Avestan, Greek, Latin, Gothic, and Slavic.

One usually distinguishes around fifty or sixty modern Armenian dialects, a number of which are now extinct. The dialects have been classified according to a phonological principle, namely the development of the consonant system, as well as a morphological one, that is the formation of the indicative present (these two features are jointly demonstrated in Table 2). Čahukyan (1972; Djahukian *apud* Greppin and Khachaturian 1986: 19–24) offers a multi-feature classification and on this basis identifies eleven dialectal groups. Armenian is written with its own script, shown in Table 1. To enhance accessibility, most of the material in this chapter is provided in a Latin-based transcription.

After a short bibliographical sketch and brief accounts of the problems of the dialectal archaisms, the classification of the Armenian dialects and the earliest dialectal isoglosses, we shall turn to the description of a number of dialects from different dialectal groups.

Table 1: The Armenian Alphabet

Arm. scr.	Translit.	Phonet.	Arm. scr.	Translit.	Phonet.
ս	a	a	ւ	m	m
բ	b	b	յ	y	j
գ	g	g	ն	n	n
դ	d	d	շ	š	ʃ
է	e	(j)e	օ	(v)o	
զ	z	z	չ'	tʰʃ	
Է	ē	ɛ	պ	p	p
Ը	ə	ə	Ջ	dʒ	
Թ	t'	th	Ր	r	r
Ժ	ž	ʒ	Տ	s	s
Ւ	i	i	Վ/Ւ	v/w	v/w
Լ	l	l	Տ	t	t
Խ	x	χ	Ր	r	r
Ծ	c	ts	Շ	c'	tʰs
Կ	k	k	Ո	u	u
Հ	h	h	Փ	p'	pʰ
Ճ	j	dʒ	Ք	k'	kʰ
Ղ	t	թ	Օ	օ	ɔ
Ճ	č	tʃ	Ֆ	f	f

2. Bibliographical sketch

The foundations of Armenian dialectology were laid by Hrač'ya Ačaryan, whose *Armenian dialectology* (1911, cf. 1909), *Armenian regional dictionary* (1913) and eleven dialect descriptions form the basic corpus of dialectological data, which are systematically included, supplemented and evaluated in his fundamental *History of the Armenian language* (AčarHLPatm 1940, 1951), *Complete grammar of the Armenian language* (AčarLiak 1955, 1957, 1971), and especially in his crowning work, the *Root dictionary of Armenian* (1971–1979).

Besides Ačaryan's works, the following general dialectological research and handbooks should be mentioned: Patkanov (1869), Yovnanean (1897), Msereanc' (1899), Łaribyan (1953), A. V. Grigoryan (1957), Jahukyan (1972), Simonyan (1979), Greppin and Khachaturian (1986), G. Mkrtč'yan (2015). Extensive phonological treatments are given in S. A. Avagyan (1973), Muradyan (1982) and Vaux (1998), as well as in a series of papers by Weitenberg. General overviews on some aspects of Armenian dialectology can be found in Weitenberg (2002) and Hovsep'yan and Geworgyan (2013). Besides the aforementioned Ačaryan (1913), important dialectological dictionaries are Ačaryan (1902), Amatuni (1912) and HayLezBrbBař 1–7 (2001–2012). The most recent description of all the Armenian dialects with references can be found in Martirosyan (2013). Dialect maps

can be found in Ačaryan (1909, 1911), Čahukyan (1972), and Hewsen (2001: 227, Nr. 218).

3. Dialectal diversity in the Classical period and earlier

The existence of dialectal diversity in the Classical period is much debated. Scholars consider a number of traces of early diversity (Ałayan 1958; Winter 1966: 205; Kortlandt 1980: 105, 2003: 32; Beekes 2003: 142–143; Clackson 2004–2005: 154). A frequently cited example is dial. **lizu* vs. Classical *lezu* ‘tongue’ (Meillet 1936: 11; Viredaz 2003: 76). The modern dialects preserve important data for the reconstruction of the oldest history of the language.

All the modern dialects have fully participated in the fixation of the proto-Armenian accent on the (prehistoric) penultimate syllable and the subsequent apocope (loss of some sounds of the final unaccented syllable). The formation of the Armenian dialects cannot thus be pushed back beyond the date of apocope. At a later stage, the accent was retracted back to the penultimate syllable in Eastern dialects.¹ It is certain, however, that Armenian dialect diversity existed in the prehistoric period (i. e. before the 5th century CE), and the modern dialects have preserved features which are not present in Classical Armenian.

Issues regarding the origin of the Armenian dialects and their existence in the classical period, as well as numerous archaic dialectal words and features are dealt with in several monographs (Ačaryan-AčarHLPatm1 1940b: 118–122; Ačaryan-AčarHLPatm2 1951: 114–141, 324–439; Čahukyan 1972, 1985, 1987: 252–256, 274–279; Simonyan 1979) and a number of papers mostly by Armenian scholars. In these studies, dialectal archaisms are often represented as a preservation of what has been lost in the corresponding classical forms. This view should be verified in each individual case. Many deviant features that have been assigned to great antiquity can, in fact, be easily explained within the framework of recent internal developments.

The deviant dialectal forms and features are systematically treated in the recent etymological dictionary of the inherited lexicon, Martirosyan (2010).

¹ See Ačaryan (1899: 184–186; AčarHLPatm2 1951: 371–372); Weitenberg (2002: 148–152).

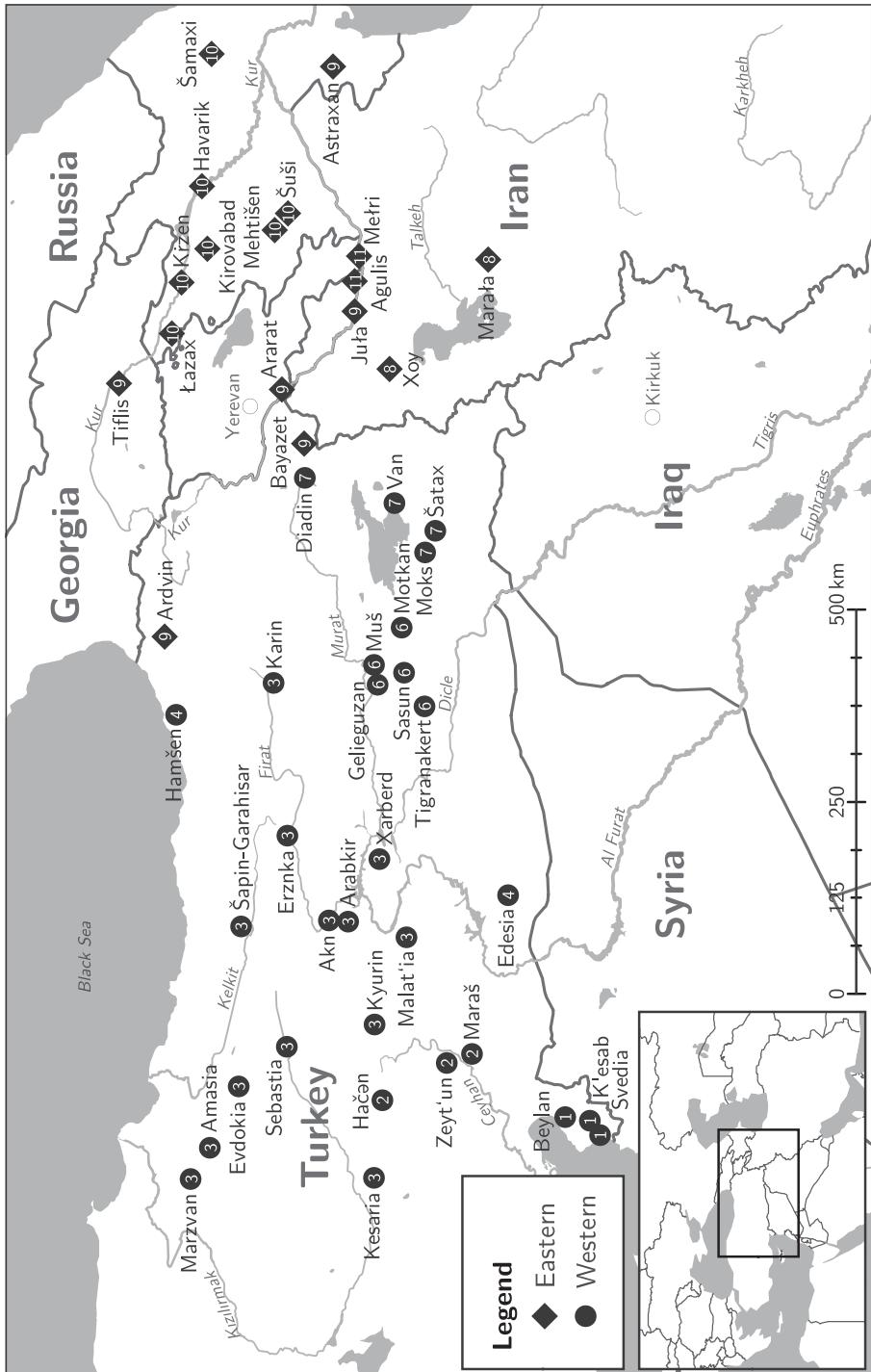


Figure 1: Location of Armenian dialects discussed in this chapter

4. Classification of Armenian dialects

4.1. The consonant system

The Armenian dialects play an important role in the discussion of two major issues in Indo-European phonology: glottals and voiced aspirates.

Classical Armenian possesses a three-fold opposition: voiced – voiceless – voiceless aspirate. Most of the dialects display varying developments of the voiced and voiceless stops, whereas the voiceless aspirate series is stable everywhere. The dialects of the Group 6 (T'iflis, Ardvin, Ararat/Lori, Agulis, Křzen, etc.) have retained the system intact. Particularly interesting are the dialects of Group 4 (Sasun, Cilicia, Svedia, etc.) which display the following steps of the so-called Second/Modern Armenian Sound Shift: voiceless > voiced; voiced > voiceless. This seems to require an interchange of two features, voiced and voiceless (cf. Class. *Tigran* > dial. *Dikran*), which is impossible from a relative chronological point of view. Therefore other features have been adduced in the discussion. On the one hand, the Classical Armenian voiced stops have been interpreted as voiced aspirates that directly continued the Proto-Indo-European series **b^h*/**d^h*/**g^h*. In these terms, the original opposition would seem to have been preserved in Group 2 (Karin, Muš, parts of Ararat, etc.). On the other hand, the glottalic character of the voiceless stops in some dialects has been interpreted as an inherited, Indo-European, feature. These two issues have been heavily debated, however. Are these phonological features archaic remnants of Proto-Armenian phonology or they are recent innovations in Armenian dialects? This basic question still awaits an answer.²

4.2. The formation of the indicative present

The morphological classification of the Armenian dialects, developed mostly by Ačaryan and Łaribyan, is based on the formation of the present indicative. In Classical Armenian, the present indicative was of a simple synthetic type, e. g. *sirem* ‘I love’, *mnam* ‘I stay’. Armenian gradually used this form to express the subjunctive and future, while a new present indicative developed in various ways: *kə* present formative particle + finite verb (*kə sirem*, *kə sires*, etc.), participle [originally dative-locative case] + copula (*sirum em*), infinitive + copula (*sirel em*), infinitive in locative case + copula (*sirelis / sire[li]s em*), as well as action noun + copula (*lalman em*, cf. CA action nouns in *-umn*, gen. *-man*). The development of these characteristic features dates to around the 11th century. Xotorjür and Aramo have preserved the classical present intact.

² For a discussion and references, see Khachaturian (1983); Pisowicz (1997); Vaux (1997, 1998: 7–12, 211–241); Weitenberg (2002: 146–151) and Haneyan (2010).

The particles *kə/ku*, *ha* and others are also involved in making the present progressive/continuous forms in western dialects. In the village of Galaduran the particle *ha(y)* forms both the simple and continuous presents, and *kə* is used for the subjunctive. The particle *kə/ku* (> *gə/gu*) is often doubled, e. g. *Zeyt'un gə g-ařnum* 'I am taking'; verbs starting with a consonant do not distinguish between the simple and progressive or continuous present (Ačaryan 2003: 246–252). A combination with *gor* (a by-form of the progressive / continuous particle) is observable in other dialects, such as Polis *g-ert'am gor* 'I am going' (Ačaryan 1941: 150–151), Sebastia *b'erem gor* 'I am bringing', *g-udeyi gor* 'I was eating' (Ačaryan 1911: 227). Some dialects combine *kə* with other particles: Aslanbek *gə sirim háye* 'I am loving' (Ačaryan 1951: 353), Trapizon *g-ařnim er* 'I was taking' (Ačaryan 1951: 344), Šapin-Garahissar *g-əsim dar* 'I am saying' (Ačaryan 1911: 174–175). Sivri-Hisar uses *kar*, *kikenan* or *er/erer* (Mkrtč'yan 1995: 212; 2006: 115, 181–184), and Nikomedia uses *yɔr* (Ačaryan 1951: 353).

Hamšen uses the verb *unim* 'to have' as a progressive marker: *b'erim uni* (or *b'erim g-uni*, with the same particle *gə*) 'I am bringing' (ClArm. 1SG.PRS *berem* 'to bring, carry'), *b'ereya g-uni* 'I was carrying' (ClArm. 1SG.IMPF *berēi*); or involves the infinitive: *yes eguš ā* 'I am coming' (ClArm. *es* 'I'; *ek-* aor. stem of the suppletive verb *gam* 'to come'; 3SG copula *ē* 'is'), *tun eguš ā* 'you are coming' (ClArm. *du* SG 'you'; *ek-* AOR stem of the suppletive verb *gam* 'to come'; 3.SG copula *ē* 'is'), etc.; or with a subject in the genitive: *imə/ims eguš ā* 'I am coming', *k'ugə(d) eguš ā* 'you are coming', with *im* 'my' and *k'o* 'your' respectively (Ačaryan 1947: 140–141). Note also Arabkir *g-ert'am (ə)nə* 'I am going' (Ačaryan 1911: 216; 1951: 349).

For the material and a discussion on the (progressive) present tense, see especially G. Geworgyan (1985, 1994, 2010). For a general overview in English, see Weitenberg (2002: 145–146).

4.3. Dialect classification based on phonology and morphology

The phonological and morphological classifications are jointly demonstrated in Table 2, which cannot, it must be said, pretend to a high degree of accuracy and completeness. The main problem here is, of course, that our data concerning numerous dialects are ambiguous or incomplete.

For the sake of simplicity only the word-initial reflexes are taken into account. There is a certain discrepancy between various descriptions of the consonant systems of some dialects, e. g. Svedia, Xarberd and Erznka. It is noteworthy that Sivri-Hisar and Yozgat'/Gamirk', according to N. Mkrtč'yan's data, display *k-k-k'* and therefore belong to the Group 7, that is Van-Łarabał, though they are very far from them, both geographically and linguistically.

In the column "Present indicative" we find the aforementioned particles and first person singular forms of the following verbs: *alam* 'to grind', *amač'em* 'to

be ashamed', *anc'anem* 'to pass', *arñum* 'to take', *asem* 'to say, tell', *arjakem* 'to untie, loosen', *banam* 'to open', *gam* 'to come', *grem* 'to write', *ert'am* 'to go', *lam* 'to weep', *lsem* 'to hear, listen', *xałam* 'to move, play', *xawsim / xōsim* 'to speak, talk', *xmem* 'to drink', *kardam* 'to shout, call, read', *mnam* 'to remain, stay', *nstim* 'to sit', *sirem* 'to love', *tam* 'to give', *uze(na)m* 'to will, wish', *utem* 'to eat'.

In the last two columns of the table one can find references to linguistic, ethnographical and folklore literature concerning the individual dialect areas. The dictionaries are given in italics.

4.4. Čahukyan's multi-feature classification

Čahukyan (1972)³ offers a multi-feature classification and on this basis identifies the following dialectal groups:

A. Western grouping

- I. The Antiok' or extreme South-Western group: K'esab/Svedia, Beylan.
- II. (I–III) The Cilicia or South-Western intergroup: Hačən, Maraš/Zeyt'un.
- III. The Asia Minor or Western group: Karin, Šapin-Garahisar, Sebastia, Evdokia, Marzvan/Amasia, Nor-Naxijewan/Crimea, Polis (Constantinople), Sivri-Hisar, Kyurin, Syolyoz, Malat'ia, Kesaria, Xarberd/Erznka, Aslanbek, Akn, Arabkir.
- IV. (I–III–VI) The Hamşen or North-Western intergroup: Hamşen, Edesia.
- V. (III–VI) The Ařtial (Transylvania) or extreme North-Western intergroup: Ařtial (Transsylvania).
- VI. The Muš/Tigranakert or South-Central group: Muš, Talvorik/Motkan, Sasun/Gelieguzan, Tigranakert.
- VII. (VI–VIII) The Van or Southern intergroup: Van, Moks, Šatax, Diadin.

B. Eastern grouping

- VIII. The Xoy/Marała or South-Eastern group: Xoy, Marała.
- IX. The Ararat or North-Eastern group
- X. Ararat, Juła, Bayazet, Astraxan, Ardvin/T'iflis.
- XI. (IX–XI) The Łarabał/Šamaxi or extreme North-Eastern intergroup: Mehtišen, Łarabał, Łazax/Kirovabad, Krzen, Havarik, Šamaxi, Burdur.
- XII. The Agulis/Mełri group: Agulis, Mełri.

³ Čahukyan (Djahukian) *apud* Greppin and Khachaturian (1986: 19–24).

Table 2: Overview of dialect classification

Indo-Eur.		*g ^h	*g	*k				Ethn. & Folklore
Classical		g(')	k	k'	Pres. indicative	Dial. deser.		
Xarberd/ Derşim	Gr. 1	g'/k	g	k'	<i>gə sirim, g-uzim</i>	Bahramyan (1960); Gasparyan (1979), Kostandyan (1982)	Andranik (1900); Sargisean (1932); Halajyan (1973); Hayk (1959); Srapean (1960)	
Şapin-Garah.	Gr. 1	g‘	g	k‘	<i>gə sirəm,</i> <i>g-ərt'əm</i>	H. S. Xač‘attyan (1985)	Tēōvlēt‘ean (1954)	
Arabkir	Gr. 1	g‘	g	k‘	<i>sirim gu,</i> <i>g-ərt'əm ga,</i> <i>gə g-ərt'əm</i>	Dawit‘-Bék (1919)	Baxtikian (1934); P’olataean (1969)	
Sebastia	Gr. 1	g‘	g	k‘	<i>sirəm gə,</i> <i>g-ərt'əm gə</i>	<i>Gabikian</i> (1952)	<i>Gabikian</i> (1952: 591–682)	
Akn	Gr. 1	g‘	g	k‘	<i>giū sirim, g-üdim</i>	Maxudianz (1911); Mak’sudeanc‘ (1910); Gabriēlean (1912)	Čanikean (1895); Azatean (1943); K’ec‘ean & Parsamean (1952)	
Ařial	Gr. 1	g‘	g	k‘	<i>gi sirim</i>	Hanusz (1887–1889); Ačaiyan (1953)		
Crimea / Nor- Naxjewan	Gr. 1	g‘	g	k‘	<i>k‘(i)-sirim, gə</i> <i>lsim</i>	Ačaiyan (1925); <i>Jałasjyan</i> (2012)	Patkanov (1875.1); P’ork’šeyan (1971)	
Hamşen	Gr. 1/4	g'/k	g	k‘	<i>sirim gu, g-udim</i>	Ačaiyan (1947); Bläsing (1992, 1995, 2007); Vaux (2000–2001, 2007)	Muradean (1901); T’orlak’yan (1981, 1986); Gurunyan (1991); Simonian (2007); JaynHamş 1–4; HamşHamş (2007); Vardanyan (2009); Hovannessian (2009)	

Indo-Eur.		*g ^h	*g	*k				
Classical		g(')	k	k'	Pres. indicative	Dial. descr.		Ethn. & Folklore
Xotorjur	Gr. 1/2	g‘	k/g	k‘	<i>sirim, mənam;</i> <i>k-asin</i>	Kostandyan (1985); Bahranyan (1976)	Gawarac'i (1903); Hačean (1907); YušamXotorj (1964)	
Erzka	Gr. 2/7	g‘/k	k	k‘	<i>k/gə sirim,</i> <i>k-udam</i>	Kostandyan (1979)	Tēr-Vardanean (1968)	
Karin	Gr. 2	g‘	k	k‘	<i>sirem k/gə,</i> <i>k-udem k/gə</i>	Tomson (1887); H. M. Mkrtč'yan (1952); Hakobyan (1974; 409–437)	Lalayean (1892, 1897, 1983); Mxit'reaneanc' (1901); Malxasyanc' (1958); Hakobyan (1974)	
Muš	Gr. 2	g‘	k	k‘	<i>kə sirim, k-atam</i>	Patkanov (1875.2); Msereanc‘ (1897–1901); Baldasaryan T‘ap‘alc‘yan (1958); Tarōnean (1961)	Sedrakean (1874); Lalayean (1917a); Melik‘ean (1964); Movsisyan (1972); Hovannissian (2001)	
Alaškert	Gr. 2	g‘	k	k‘	<i>kə sirim, k-atam</i>	Madat‘yan (1985)	Yovsēp‘eanc' (1892: 47); Nždehean (1899–1910)	
Ararat	Gr. 2	g‘	k	k‘	<i>sirə/uum em</i>	Nawasardeanc' (1903); Kapançjan (1975); Markosyan (1989)		
Nor Juła	Gr. 2	g‘	k	k‘	<i>sirum am,</i> <i>mořanum am;</i> <i>lalman am,</i> <i>talman am</i>	Ačaryan (1940a)	Eremean (1923, 1930)	
Polis	Gr. 3	g	g	k‘	<i>gə sirəm,</i> <i>g-er'i'am</i>	Ačaryan (1941)	Svazlyan (2000)	

Indo-Eur.		*g ^h	*g	*k			Ethn. & Folklore
Classical	g(‘)	k	k‘	Pres. indicative	Dial. deser.		
Trapizon	Gr. 3 g	g	k‘	sirim <i>gu</i> , <i>g-užim</i>	Ačaiyan (1911: 178)	Yovakimean (1967); T'orilak'y.	
						(1986); Hovannissian (2009)	
Evdokia	Gr. 3 g	g	k‘	<i>gə sirim</i> ,	Gazančean (1899)	Alpöyačean (1952)	
				<i>g-er'i'am</i>			
Marzvan	Gr. 3 g	g	k‘	<i>gə sirim</i> , <i>gə banam</i>	T'umayean (1930)		
Amasia	Gr. 3 g	g	k‘	(<i>instim ga</i> , <i>g-er'i'am ga</i>) [3SG <i>karda-v!</i>]	Ačaiyan (1951: 350)		
Kesaria	Gr. 3 g	g	k‘	<i>ga sirəm</i> , <i>ga g-arium</i>	Ant'osyan (1961)	Grigorean & Garakēōzean (1963)	
St'anoz	Gr. 3 g	g	k‘	<i>g-er'i'am kar</i>	N.A. Mkrtč'yan (2006: 202–222, 293–294)	Ötean-Gashbarean (1968)	
Aynt'ap	Gr. 3 g	g	k‘	—	Vaux (1999–2000)	Sarafean (1953)	
Sasun	Gr. 4 k	g	k‘	<i>gə sirəm</i> , <i>g-er'i'am</i>	Petoyan (1954)	K' alant'ar (1895); Tomaschek (1896); Karapetyan (1962); Petoyan (1965); R. Xač'atryan (1999)	
Zeyt'un	Gr. 4 k/g‘	g	k‘	<i>gə siyim</i> , <i>g-iy't'm</i>	Mak'sudeanc' (1911), Ačaiyan (2003)	Allahērēcean (1884); Galustean (1934); Guyumčean (1990)	
Hačən	Gr. 4 k/g‘	g	k‘	<i>gə siyem</i> , <i>gə gərñum</i>	Ačaiyan (2003); Gasparyan (1966)	Pōlosean (1942)	

Indo-Eur.		*g ^h	*g	*k			
Classical		g(‘)	k	k‘	Pres. indicative	Dial. descr.	Ethn. & Folklore
Svedia	Gr. 4	k/g‘	g	k‘	gə <i>sirim</i> , g’-ärnim	Andreasyan (1967); Ačaiyan (2003); Hananyan (1995)	Svazlyan (1984, 1994); Gyozalyan (2001)
Aramo	Gr. 4	k	g	k‘	(hay) <i>sireym</i> , (hay) <i>urt’um</i>	Ayceamm (1907); Łaribyan (1958: 9–77)	
K’abusie	Gr. 4	k	g	k‘	gu <i>sirym</i> , geu <i>urt’um</i>	Łaribyan (1958: 78)	
K’esab	Gr. 4	k	g	k‘	ha <i>sirym</i> , hay <i>urt’um</i>	Łaribyan (1955); Č’olak’ean (1986)	
Beylan	Gr. 4	k	g	k‘	gä <i>sirym</i> , gä <i>garıom</i>	Łaribyan (1955)	
Malat’ia	Gr. 5	k‘	g	k‘	<i>sirim a</i> , g-erīt’am <i>a</i>	Danielyan (1967)	Alpoyačeān (1961)
Nikomedie/ Aslanbek	Gr. 5	k‘/g‘	g	k‘	gə <i>sirym</i> , g-iäm (contin. + <i>hayε</i>)	Ačaiyan (1898a, 1898b); Vaux (2001)	Tēr-Yakobean (1960)
Edesia/ Urha	Gr. 5	k‘	g	k‘	<i>sirem a</i> / gə <i>sirem</i> g-erīt’am a	Łaribyan (1958: 146); Haneyan (1982)	
Tigranakert	Gr. 5	k‘	g	k‘	gə <i>sirym</i> , g-iuzim	Haneyan (1978)	
Tiflis	Gr. 6	g	k	k‘	sirum im , ku <i>sirim</i>	Petermann (1867b); Tomson (1890)	Tēr-Atēk’sandrean (1885); Alababean (1901)

Indo-Eur.		*g ^h	*g	*k				Ethn. & Folklore
Classical	g(‘)	k	k‘	Pres. indicative	Dial. deser.			
Ararat/ Lori	Gr. 6 g	k	k‘	sírəm em, mnəm em	Asatryan (1968)	T. Geworgyan (1999); Lalayean (1903, 2004)		
Aqulis	Gr. 6 g	k	k‘	sáyriñm em, n-đam em	Petermann (1867a); Sargseanc‘ (1883); Ačaiyan (1935); C’ha təsnäs em	Ayyazyan (1984) Zak’aryan (2008)		
Ares/ Havarik	Gr. 6 g	k	k‘	1SG sirell̄s em, sirelli yem 2SG sirel̄m es 3SG sirel̄lin i	Tēr-Pōtosean (1921– 1922); Lusenc‘ (1982)			
Křzen	Gr. 6 g	k	k‘	atlaš em vs. em atali	Badranyan (1961)			
Ardvin	Gr. 6 g	k	k‘	sirelis im, mnalis im	Ačaiyan (1911; 291–292; 1926: 267); Łaribyan (1939: 33–49; 1953: 281–292, 329–342), Alaverdyan (1968ab).			
Laradat	Gr. 6 g	k	k‘	sirilis em, mnalis em	Łaribyan (1953: 281– 304)	H. Hovsep’yan (2009, 1–2)		
D(i)zmar	Gr. 6 g	k	k‘	sirijs em, mnayis em	Łaribyan (1953: 281– 292)	H. Hovsep’yan (2009, 1: 60)		
Medri	Gr. 6 g	k	k‘	siris im, menas im	Łaribyan (1939: 83–106); Alayan (1954)			

Indo-Eur.		*g ^h	*g	*k			
Classical	g(‘)	k	k'	Pres. indicative	Dial. descr.	Ethn. & Folklore	
Kar'ewan	Gr. 6 g	k	k'	siriy im / im sirts	Łaribyan (1939: 59–82); Muradyan (1960)		
Kak'avaberd	Gr. 6 g/k	k	k'	xəmūm em, ásəm em; A. məndpis im B. məndis im	Muradyan (1967)		
Astraxan	Gr. 7 k	k	k'	asəm em, lsəm em	aćaiyan (1911: 82–86); Aćaiyan (1951: 333)	G. G. Geworgyan (1980); Łaziyān & Vardanyan (2004)	
Goris/ Syunik‘	Gr. 7 k	k	k'	sirum em, xəmum em	Vartapetjan (1962); Margaryan (1975)	Lisic'yan (1969)	
Łarabat	Gr. 7 k	k	k'	síris məris məsərx	aćaiyan (1899); Davtyan (1966); Sargsyan (2013)	Barxudareanc‘ (1883); Barxutareanc‘ (1883); Melik‘- Şahnazareanc‘ (1907 f); Lisic'yan (1981); Łaziyān (1983); Lalayean (1897a, 1988); Grigoryan- Spandaryan (1971); L. Harut'yunyan (1991)	
Tavuş	Gr. 7 k	k	k'	sirum em, mnem em	Mežunc‘ (1989)	Xemčyan (2000, 2008)	
Burdur	Gr. 7 k	k	k'	sirum em, mnun em	Mkrtč'yan (1971, 2006: 304)		
Şamaxi	Gr. 7 k	k	k'	A. sirüm äm, mənam am, B. k'iris em, mənays em	Badranyan (1964)	Łaziyān & Vardanyan (2004)	

Indo-Eur.		*g ^h	*g	*k	Pres. indicative	Dial. deser.	Ethn. & Folklore
Classical	g(‘)	k	k‘				
Šatax	Gr. 7	k	k	k‘	<i>kʰərijs əm,</i> <i>mənayis əm</i>	Łaribyan (1939: 50–58); Davtyan (1966)	
Hadrut‘	Gr. 7	k	k	k‘	<i>siris əm, mənas</i> <i>əm</i>	Łaribyan (1939: 107–127); Połosyan (1965); Davtyan (1966)	
Urmia	Gr. 7	k	k	k‘	<i>kʰires em/s, 3SG</i> <i>kʰirel i;</i> <i>mənas em/s, 3SG</i> <i>mənal i</i>	Łaribyan (1939: 128–146; 1953: 342–358); Asatryan (1962)	
Xoy	Gr. 7	k	k	k‘	<i>üzel em, et əl em</i>	Asatryan (1962)	
Marala	Gr. 7	k	k	k‘	<i>üzeli im,</i> <i>amč’əmalð yim</i>	Ačaiyan (1926)	
Hin Ĵula	Gr. 7	k	k	k‘	<i>talmān am</i>	Ačaiyan (1940a: 27–32)	
Van	Gr. 7	k	k	k‘	<i>kə sīrem,</i> <i>k-uzem,</i> <i>ku kəñam</i>	Ačaiyan (1952); <i>Turşyan</i> (2018)	Sedrakean (1874); Tēr-Sargsenc‘ (1875); Šērenc‘ (1885–1999); Lalayean (1910, 1914, 1915, 1917b); Srujaniteanc‘ (1978); S. M. Avagyan (1978); Hovannissian (2000)
Moks, Šatax	Gr. 7	k	k	k‘	<i>kə sirim,</i> <i>kə xatam,</i> <i>k-ənc’ nim</i>	M. H. Muradyan (1962, 1982); <i>Orbeli</i> (2002)	Yovsep’eanc‘ (1892); Šahpazean (1913); Lalayean 2 (1914)

Indo-Eur.		*g ^h	*g	*k			
Classical		g(')	k	k'	Pres. indicative	Dial. descr.	Ethn. & Folklore
Ozim	Gr. 7/2	k/g [‘]	k	k'	kə kar̥tām, k-arc'kim	N. E. Hovsep'yan (1966); Arewikyan (1967)	
Sivri-Hisar	Gr. 7	k	k	k'	(cont.) k-εt'am <i>er / kar,</i> <i>xɔsim kənam</i>	N. A. Mkrtč'yan (1995, 2006)	PtmSivHisHay (1965)
Yozgat' / Gamirk'	Gr. 7	k	k	k'	sirim ka	N. A. Mkrtč'yan (2006)	T' emurč'yan (1970)

5. Earliest dialectal isoglosses

5.1. Weitenberg's isoglosses

In a series of articles Weitenberg (1986, 1992, 1993, 1996, 2001) discussed a relatively homogenous set of dialectal isoglosses within a chronological framework starting from pre-literary times, all of which are anterior to the sound shift $h > x$ in the Van-Urmia group and the devoicing ($b, d, g > p, t, k$) which are dated to approximately the 7th century or later:

- Retraction of the accent to the penultima.
- Diphthongization of stressed *o* (and perhaps *e*) in initial position; all dialects diphthongize any *o* in absolute anlaut, whereas the dialects having accent retraction (Łarabał, etc.) and the dialects of the Van-Urmia region do so only in monosyllables (these are marked in Table 3 “no”).
- Monophthongization of *aw*; if *aw* (under penultimate accent) is followed by a dental stop or affricate, the *w* yields *x*, e. g. *eawt'n* ‘seven’ > *Mełri ɔxtə* (these are marked in Table 2 “yes”).
- Ačafyan's Law (fronting of back vowels after voiced obstruents).
- Devoicing $b, d, g > p, t, k$ (7th century or later).

In Weitenberg (1996: 111) these isoglosses are combined in a table in relative chronological order, from left to right. In addition, the isogloss of present verbs in *-nul* vs. ClArm. *-anem* is represented by the example of *tesanem* ‘to see’.

5.2. Miscellaneous

There are more of such isoglosses, e. g. the additional *-n* mostly represented in the dialects of the old Arc‘ax-Siwnik‘ area, which is a feature originating in the pre-literary period (H. D. Muradyan 1982: 322–326; Weitenberg 1985) and monophthongization of the final stressed *-ay*, which can be traced back to the fifth century (Weitenberg 1999–2000).

5.3. Extending Weitenberg's classification

In Table 3, I present a revised and supplemented version of Weitenberg's aforementioned table in combination with data from my Table 2. The archaic set of isoglosses demonstrates an opposition between the South-Eastern periphery (Łarabał/Agulis area) on the one hand, and the Central and Western regions on the other. As Weitenberg (1996: 112) points out, the Van-Urmia intermediate area may originally have formed part of the South-Eastern area. Note also the sharp contrast between Muš and Łarabał.

Table 3: Extended classification based on Weitenberg

<i>Dialect</i>	<i>Phonological classification</i>	<i>Present indicative</i>	<i>Accent retraction</i>	<i>Diphthong o-</i>	<i>ʒxt '7'</i>	<i>Āčar̄. Law</i>	<i>Devoicing</i>	<i>tesanel</i>
<i>Ariyal</i>	<i>Gr. I</i>	<i>gi sirim</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>	<i>-nul</i>	
<i>Arabkir</i>	<i>Gr. I</i>	<i>sirim gu,</i> <i>g-ert'am ga,</i> <i>gɔ g-ert'am</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>-nal</i>	
<i>Erz̄nka</i>	<i>Gr. I/4</i>	<i>k/gɔ sirim,</i> <i>k-udam</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>-nal</i>	
<i>Hamšen</i>	<i>Gr. I/4</i>	<i>sirim gu,</i> <i>g-udim</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>-nul</i>	
<i>Karin</i>	<i>Gr. 2</i>	<i>sirem k/gɔ,</i> <i>k-udem k/gɔ</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>	<i>-nul</i>	
<i>Muš</i>	<i>Gr. 2</i>	<i>kə sirim,</i> <i>k-ətəm</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>-nul</i>	
<i>Ararat</i>	<i>Gr. 2</i>	<i>sirə/u'm em</i>	<i>(yes)</i>	<i>(yes)</i>	<i>yes</i>	<i>no</i>	<i>(-nul)</i>	
<i>Nor Ĵuta</i>	<i>Gr. 2</i>	<i>sirim am,</i> <i>morānum am;</i> <i>talman am,</i> <i>talman am</i>	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>	<i>-nel</i>	
<i>Polis</i>	<i>Gr. 3</i>	<i>gɔ sirem,</i> <i>g-ert'am</i>	<i>no</i>	<i>ABSENT</i>	<i>yes</i>	<i>no</i>	<i>-nul</i>	
<i>Hač̄ən</i>	<i>Gr. 4</i>	<i>gɔ siyem, gɔ</i> <i>g-ert'am</i> <i>g-ərnum</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>-nul</i>	

<i>Dialect</i>	<i>Phonological classification</i>	<i>Present indicative</i>	<i>Accent retraction</i>	<i>Diphthong o-</i>	<i>xt '7'</i>	<i>Ačar. Law</i>	<i>Devoicing</i>	<i>tesanel</i>
<i>Svedia/ Urhā</i>	<i>Gr. 4</i> <i>gə sirim, gʷärnim</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes?</i>	<i>yes</i>	<i>-mul</i>	
<i>Edesia/ gert'am a</i>	<i>Gr. 5</i> <i>sirem a / ga sirem g-ert'am a</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>-nal</i>	
<i>Tigranakert</i>	<i>Gr. 5</i> <i>gə sirim, g-užim</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>-nal</i>	
<i>Ararat/ Lori</i>	<i>Gr. 6</i> <i>sirem em, mnen em</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>no</i>	<i>-nal</i>	
<i>Agulis</i>	<i>Gr. 6</i> <i>sávṛäm ḡam, n-ǵtam ḡam</i>	<i>yes</i>	<i>(yes)</i>	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>-nil</i>	
<i>Areš/ Havarik</i>	<i>Gr. 6</i> <i>1SG sirelis ḡam, strelj yám 2SG sireləm ḡas</i>	<i>no</i>		<i>yes</i>	<i>yes</i>	<i>no</i>	<i>-nel / -nal</i>	
<i>Metri</i>	<i>Gr. 6</i> <i>siriis im, mənas im</i>	<i>yes</i>	<i>(yes)</i>	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>-nil</i>	
<i>Karčevan</i>	<i>Gr. 6</i> <i>stiriy im / im střis</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>-(n)ul</i>	

Dialect	Phonological classification	Present indicative	Accent retraction	Diphthong o-	<i>ext 7'</i>	<i>Ačar. Law</i>	Devoicing	<i>tesanel</i>
<i>Kak'avaberd</i>	Gr. 6	xəmum em, ásəm em; <i>mənáyis im,</i> <i>-lis im</i>	yes	yes	yes	yes	no/yes	-(n)il
<i>Larabat</i>	Gr. 7	sírem em, wəməxç	yes	yes	yes	yes	yes	-nal
<i>Urmia</i>	Gr. 7	k'iress em/s, 3SG k'irel i; <i>mənas em/S,</i> <i>3SG mənal i</i>	no	yes	yes	yes	yes	-nel
<i>Marata</i>	Gr. 7	üzeli im, amč əndəj yim	no	yes	yes	yes	yes	-nel
<i>Van</i>	Gr. 7	kə sirəm, k-uzem, ku k'äm	no	yes	no	yes	yes	-nal
<i>Moks, Šatax</i>	Gr. 7	kə sirim, kə xalam, k-ənc 'nim	no	yes	no	yes	yes	-nel / -nil

6. Dialectal groups

The Armenian dialects are presented in eleven groups in accordance with Jahukyan's multi-feature classification. The dialects of the Eastern grouping (groups VIII–XI) have developed and are still spoken in the Republic of Armenia, Arc'ax (Karabagh / Լարաբալ), as well as Parskahayk'/Persian Armenia and surroundings (North West Iran). By contrast, the western dialects (groups I–VII) which had developed on the territory of the western part of the historical Armenia (that is, the eastern half of present-day Turkey) for many centuries, with a few exceptions, are no longer spoken in their original location. This is because of Turkish constant genocidal policy and especially the Armenian Genocide starting in 1915 which resulted in extermination of more than one million Armenians and constituted an amputation of Armenian life, and therefore also an amputation of Armenian dialects. Some surviving groups of the Armenian population were dispersed around the world. Their descendants have partially preserved their dialects in the Republic of Armenia, Javaxk' (now in Georgia), Russia and other countries.

For these reasons our information on the western dialects is uneven. Many of them are now almost or completely extinct. Some of them have been described before the Genocide, others are known by some secondary materials recorded from the refugees, and for the rest we practically have no information.⁴

In this survey of the dialects I shall concentrate on the most characteristic features of a number of dialects from different groups. When discussing them, I often present comparative data from dialects of other groups. The data concerning the phonological and morphological classifications and the aforementioned earliest isoglosses can be found in Tables 2 and 3 and will be specified here only when necessary.

A. Western grouping

6.1. Svedia (the Antioik' or extreme South-Western group)

Armenians have been living in the northern parts of Syria and Mesopotamia presumably since the time of Tigran the Great (1st century BCE). There is information on migration of some Armenian population to Syria in the 6th century CE. The Armenian presence in these areas became more significant especially since the 10th century. Here the Armenians called themselves *k'ist'inə* / *k'isdine*, which reflects the word *k'rystoneay* 'Christian', and the dialect, *k'ist'inəg* / *k'isding* / *k'ist'inüg* 'language of a Christian, Armenian'.

The Armenian dialects of Syria and Cilicia are very divergent and extremely interesting. Svedia is the best studied one among them. It was spoken in the Armenian villages of the Svedia district, which is located at the foot of the mountain

⁴ For a general outline on these issues and the linguistic situation in the Republic of Armenia, see Weitenberg (2002: 141–143 and 2006).

Musaleř/Musa-Dagh. There are three monographs on the dialect of Svedia: Ačaryan (2003), Andreasyan (1967) and Hananyan (1995), based on the Svedian subdialects of Hajî-Habibî, Yołun-Öluk and Xtrbek, respectively. K'esab is described in Č'olak'ean (1986). The other (sub)dialects have been poorly studied, mainly by Łaribyan.

Phonology

According to Ačaryan (2003: 401), Classical Armenian voiced stops are reflected in Svedia as voiced aspirates *b' g' d'* in initial and intervocalic positions and after *t* and *r*. In the descriptions of other subdialects, Andreasyan (1967) and Hananyan (1995), *b d g* yielded *p t k*. Table 4 provides a simplified table of the development of CA plosives into the dialect of Svedia.

Table 4: Reflex of CA plosives in Svedia

CA	S v e d i a		
	initial	internal	final
<i>b/g/d</i>	<i>b'g'd'</i>	<i>b'g'd', (N)b/g/d</i>	<i>b'g'd', (N)b/g/d</i>
<i>p/k/t</i>	<i>b/g/d</i>	<i>b/g/d</i>	<i>b/g/d</i>
<i>p'k't'</i>	<i>p'k't'</i>	<i>p'k't'</i>	<i>p'k't'</i>

One of the most characteristic phonological features of Svedia is the change *a > u* in a stressed position, where Beylan and the Cilician dialects have *ɔ*. As Ačaryan (1951: 364–378, 2003: 349, cf. 10–11) points out, the development *a > ɔ > u* logically ends in Syria, which marks the last habitat in the extreme SW corner of Armenian-speaking areas.

In Agulis, an eastern dialect that is very far from Svedia both geographically and linguistically, we find a similar change only in monosyllables. In Hamšen and Akn we find *a > ɔ* only before a nasal. Table 5 provides examples from these dialects, with the addition of some dialects that do not display *a > ɔ*, namely Ararat (an eastern dialect) and Moks/Šatax (a western dialect).

Table 5: Reflexes of CA /a/ in selected dialects

CA	gloss	Ararat	Moks/Šat.	Agulis	Hamšen	Zeyt‘un	Svedia
<i>at</i>	salt	<i>at</i>	<i>at</i>	<i>at</i>	<i>at</i>	<i>ɔt</i>	<i>ütl</i>
<i>arj</i>	bear	<i>arč‘</i>	<i>ar/rč</i>	<i>ɔrj</i>	<i>ayč‘</i>	<i>ärj‘</i>	<i>urj‘</i>
<i>art</i>	cornfield	<i>art/d</i>	<i>art</i>	<i>ɔrt</i>	<i>ayd</i>	<i>ɔyd</i>	[uɔrt] ⁵
<i>bak</i>	courtyard	<i>bag</i>	<i>päk</i>	<i>bɔk</i>	—	<i>b'ɔg</i>	<i>b'iüg</i>
<i>ban</i>	thing	<i>b'an</i>	<i>pän</i>	<i>bün</i>	<i>pɔn</i>	<i>b'ɔn</i>	<i>b'un</i>
<i>bac‘</i>	open	<i>b'ac‘</i>	<i>päc‘</i>	<i>böc‘</i>	<i>pac‘</i>	<i>b'ɔc‘</i>	<i>b'uc‘</i>
<i>lar</i>	cord	<i>lar</i>	<i>lar</i>	—	<i>lar</i>	<i>lɔy</i>	<i>lur</i>
<i>car̥</i>	tree	<i>car̥</i>	<i>car̥</i>	<i>cɔr̥</i>	<i>jar̥</i>	<i>jɔr̥</i>	<i>jur̥</i>
<i>hac‘</i>	bread	<i>hac‘</i>	<i>xac‘</i>	<i>hɔc‘</i>	<i>hac‘</i>	<i>hɔc‘</i>	<i>huc‘</i>
<i>maz</i>	hair	<i>maz</i>	<i>maz</i>	<i>mɔz</i>	<i>maz</i>	<i>mɔz</i>	<i>muz</i>
<i>tanj</i>	pear	<i>tanj</i>	<i>tanj</i>	<i>tunj</i>	<i>dɔnj</i>	<i>dɔnj</i>	<i>dunj</i>
<i>tasn</i>	ten	<i>tassə</i>	<i>tas</i>	<i>tóssə</i>	<i>dasə</i>	<i>dɔsə</i>	<i>dussə</i>
<i>k‘ar</i>	stone	<i>k‘ar</i>	<i>k‘ar</i>	<i>k‘ɔr</i>	<i>k‘ar</i>	<i>k‘ɔy</i>	<i>k‘ur</i>
<i>agraw</i>	crow	<i>akrav</i>	[ak‘rav] ⁶	<i>ákraV</i>	—	<i>ag‘rɔv</i>	<i>äg‘ruv</i>
<i>amar̥n</i>	summer	<i>amar̥</i>	<i>amar̥</i>	<i>ámaṛ(nə)</i>	<i>ɔmaṛ</i>	<i>amɔṛ</i>	<i>amuṛ</i>
<i>ardar</i>	righteous	<i>art‘ar</i>	<i>artär</i>	<i>árt‘ar</i>	<i>art‘tar</i>	<i>ayd‘ɔy</i>	<i>ard‘ur</i>
<i>arcat‘</i>	silver	<i>arčat‘</i>	<i>arčat‘</i>	<i>árcat‘</i>	<i>ayyat‘</i>	<i>ayjɔt‘</i>	<i>arjut‘</i>
<i>danak</i>	knife	<i>d‘anak</i>	<i>tänak</i>	<i>dánäk</i>	<i>tɔnag</i>	<i>d‘anɔg</i>	<i>d‘änüg</i>
<i>mangat̥</i>	sickle	<i>mangał</i>	<i>mangvät̥</i>	<i>mangvät̥</i>	<i>mɔngat̥</i>	<i>mangɔł</i>	<i>mängüt̥</i>
<i>pařaw</i>	old woman	<i>pařav</i>	<i>pařav</i>	<i>pářav</i>	<i>bařav</i>	<i>bařɔb</i>	<i>bařuv</i>
<i>gawazan</i>	stick	<i>g‘avazan</i>	<i>k‘äväzan</i>	—	<i>kavazɔn</i>	<i>g‘avɔn</i>	<i>g‘äväzun</i>
<i>hnazand</i>	obedient	<i>hnazand</i>	<i>xnazänd</i>	<i>hnázand</i>	—	<i>hnazɔnd‘</i>	<i>hnazund‘</i>
<i>patuhas</i>	punishment	—	<i>patuxas</i>	—	—	<i>baduhɔs</i>	<i>bädihus</i>

⁵ This is the K‘esab form (Č‘olak‘ean 1986: 197a). The word is not recorded in Svedia itself.

⁶ Cf. HAB [Hayerēn armatakan bařaran] vol. 1: 80a.

Nominal morphology

An important morphological peculiarity is the instrumental marker *-um* in Svedia, which is not found in other Armenian (sub)dialects of Syria. K'esab has *-uən* (Č'olak'ean 1986: 68); K'abusie has *-u* for inanimate objects and /genitive + *hid* 'with/' for people (Łaribyan 1958: 96).

For K'abusie, Łaribyan (1958: 78–79, 83, 96–99) records a remarkable contrast: nominative-accusative indefinite *murt* 'vs. definite *ma"rt*', cf. *äs ma"rt* 'this man'; the preposition *z-* in the accusative, as in CA, is found in a definite context. Note also the preposition *i* / *y-* in an allative usage: *is i dün g-ürt um* 'I am going home'; *y-ambo"r g-ürt um* 'I am going to the barn'.

Table 6 provides the paradigms of *hac* 'i-stem 'bread', and *mard*, *o*-stem 'man, human', in Svedia, Zeyt'un, (Ačaryan 2003: 165, 184, 460–461) and K'abusie (Łaribyan 1958: 98). In the following table N = nominative, ACC = accusative, GD = genitive/dative, ABL = ablative, I = instrumental:

Table 6: Nominal paradigms in selected dialects

		CA	Zeyt'un	Svedia (indef. and def.)	
SG	N	<i>hac</i> '	<i>hac</i> '	<i>huc</i> '	<i>huc</i> 'ə
	ACC	(<i>z</i>) <i>hac</i> '	<i>hac</i> '	<i>huc</i> '	<i>huc</i> 'ə
	GD	<i>hac</i> '-i	<i>hac</i> '-ə	<i>häc</i> 'c'-ə	<i>häc</i> 'c'en
	ABL	<i>i hac</i> '-ē	<i>hac</i> '-e	<i>häc</i> 'c'-ε	<i>häc</i> 'c'in
	I	<i>hac</i> '-i-w	<i>hac</i> '-öv	<i>häc</i> 'c'-um	<i>häc</i> 'c'umə
PL	N	<i>hac</i> '-k'	<i>hac</i> '-iy	<i>häc</i> 'c'-ir	<i>häc</i> 'c'irə
	ACC	(<i>z</i>) <i>hac</i> '-s	<i>hac</i> '-iy	<i>häc</i> 'c'-ir	<i>häc</i> 'c'irə
	GD	<i>hac</i> '-i-c'	<i>hac</i> '-iy-u	<i>häc</i> 'c'-ira	<i>häc</i> 'c'iren
	ABL	<i>i hac</i> '-i-c'	<i>hac</i> '-iy-e	<i>häc</i> 'c'-ire	<i>häc</i> 'c'irin
	I	<i>hac</i> '-i-w-k'	<i>hac</i> '-iy-öv	<i>häc</i> 'c'irum	<i>häc</i> 'c'irum
SG		CA	Zeyt'un	K'abusie	Svedia
	N	<i>mard</i>	<i>mayd</i> '	<i>murt</i> '	<i>muṛd</i> '
	ACC	(<i>z</i>) <i>mard</i>	<i>mayd</i> '	<i>z-ma"rt</i> '	<i>muṛd</i> '
	GD	<i>mard</i> -oy	<i>mayd</i> '-o	<i>mart</i> '-ən	<i>maṛd</i> '-ə
	ABL	<i>i mard</i> -oy	<i>mayd</i> '-e	G + <i>irmeyn</i>	<i>maṛd</i> '-ε
	I	<i>mard</i> -o-v	<i>mayd</i> '-öv	G + <i>heyd</i>	<i>maṛd</i> '-um

Verbal morphology

Table 7 shows the present and imperfect paradigms of *sirem* ‘to love’ in the Haji-Habibli subdialect (Ačaryan 2003: 482–490) and *grem* ‘to write’ in the Xtrbek subdialect (Hananyan 1995: 125–126).

Table 7: Present and imperfect paradigms, *sirem* ‘to love’ & *grem* ‘to write’ in Svedia (Haji-Habibli subdialect)

Present indicative			Imperfect		
	CA	Svedia 1	CA	Svedia 1	Svedia 2
SG	<i>sirem</i> , <i>grem</i>	<i>gə sirim</i>	<i>geu kərim</i>	<i>sirēi</i> , <i>grēi</i>	<i>gə sirer ε</i>
	<i>sires</i> , <i>gres</i>	<i>gə siris</i>	<i>geu kəris</i>	<i>sirēir</i> , <i>grēir</i>	<i>gə sirer</i>
	<i>sirē</i> , <i>grē</i>	<i>gə sire</i>	<i>geu kəri</i>	<i>sirēr</i> , <i>grēr</i>	<i>gə sirir</i>
PL	<i>siremk'</i> , <i>gremk'</i>	<i>gə sirink'</i>	<i>geu kərənk</i>	<i>sirēak'</i> , <i>grēak'</i>	<i>gə sirer ink'</i>
	<i>sirēk'</i> , <i>grēk'</i>	<i>gə sirik'</i>	<i>geu kərək</i>	<i>sirēik'</i> , <i>grēik'</i>	<i>gə sirer ik'</i>
	<i>siren</i> , <i>gren</i>	<i>gə sirin</i>	<i>geu kərin</i>	<i>sirēin</i> , <i>grēin</i>	<i>gə sirer in</i>

According to Łaribyan (1953: 444–445; 1955: 196, 201–202) and Č'olak'ean (1986: 122), in K'esab the present indicative is formed with the particle *ha* or *hay*. In the village of Galaduran this particle forms both the simple and continuous presents. The particle *kə* is used for the subjunctive. The imperfect is composed of a participle in *-er* and the copula. The latter is dropped in the third person singular form. Table 8 gives the paradigms of *kardam* ‘to shout, read’ and *grem* ‘to write’.

Table 8: Present and imperfect verbs in the Galaduran dialect

Present				
	CA	K'esab	CA	K'esab
SG	<i>kardam</i>	<i>ha gartum</i>	<i>grem</i>	<i>ha krim</i>
	<i>kardas</i>	<i>ha gartu's</i>	<i>gres</i>	<i>ha kris</i>
	<i>karday</i>	<i>ha gartu</i>	<i>grē</i>	<i>ha kri</i>
PL	<i>kardamk'</i>	<i>ha gartunk'</i>	<i>gremk'</i>	<i>ha krink'</i>
	<i>kardayk'</i>	<i>ha gartäk'</i>	<i>grēk'</i>	<i>ha krik'</i>
	<i>kardan</i>	<i>ha gartun</i>	<i>gren</i>	<i>ha krin</i>
Imperfect				
SG	<i>kardayi</i>	<i>ha garter em</i>	<i>grēi</i>	<i>ha krer εm</i>
	<i>kardayir</i>	<i>ha garter es</i>	<i>grēir</i>	<i>ha krer εs</i>
	<i>kardayr</i>	<i>ha garter</i>	<i>grēr</i>	<i>ha krer</i>
PL	<i>kardayak'</i>	<i>ha garter enk'</i>	<i>grēak'</i>	<i>ha krer enk'</i>
	<i>kardayik'</i>	<i>ha garter ek'</i>	<i>grēik'</i>	<i>ha krer ek'</i>
	<i>kardayin</i>	<i>ha garter en</i>	<i>grēin</i>	<i>ha krer en</i>

Unlike the neighbouring K‘esab, the Beylan dialect forms the indicative present and imperfect with the particle *gä* and therefore belongs to the *kə*-group. The imperfect is also marked by a postposed particle *di* throughout the paradigm excluding the third- singular form.⁷ Table 9 provides the paradigms of *kardam* ‘to shout, read’ and *grem* ‘to write’.

Table 9: Present and imperfect forms in the Beylan dialect

Present						
	CA	Beylan	CA	Beylan	CA	Beylan
SG	<i>kardam</i>	<i>gä gartəm</i>	<i>grem</i>	<i>gä kərim</i>	<i>tesanem</i>	<i>gä disnum</i>
	<i>kardas</i>	<i>gä gartəs</i>	<i>gres</i>	<i>gä kəris</i>	<i>tesanes</i>	<i>gä disnus</i>
	<i>karday</i>	<i>gä gartə</i>	<i>grē</i>	<i>gä kəre</i>	<i>tesanē</i>	<i>gä disnu</i>
PL	<i>kardamk'</i>	<i>gä gartənk'</i>	<i>gremk'</i>	<i>gä kərink'</i>	<i>tesanemk'</i>	<i>gä disnunk'</i>
	<i>kardayk'</i>	<i>gä gartek'</i>	<i>grēk'</i>	<i>gä kərik'</i>	<i>tesanēk'</i>	<i>gä disnuk'</i>
	<i>kardan</i>	<i>gä gartən</i>	<i>gren</i>	<i>gä kərin</i>	<i>tesanen</i>	<i>gä disnun</i>
Imperfect						
SG	<i>kardayi</i>	<i>gä garti di</i>	<i>grēi</i>	<i>gä kəri di</i>	<i>tesanēi</i>	<i>gä disni di</i>
	<i>kardayir</i>	<i>gä garter di</i>	<i>grēir</i>	<i>gä kərir di</i>	<i>tesanēir</i>	<i>gä disnir di</i>
	<i>kardayr</i>	<i>gä garter</i>	<i>grēr</i>	<i>gä kərer</i>	<i>tesanēr</i>	<i>gä disner</i>
PL	<i>kardayak'</i>	<i>gä gartink' di</i>	<i>grēak'</i>	<i>gä kərink' di</i>	<i>tesanēak'</i>	<i>gä disnink' di</i>
	<i>kardayik'</i>	<i>gä gartik' di</i>	<i>grēik'</i>	<i>gä kərik' di</i>	<i>tesanēik'</i>	<i>gä disnik' di</i>
	<i>kardayin</i>	<i>gä gartin di</i>	<i>grēin</i>	<i>gä kərin di</i>	<i>tesanēin</i>	<i>gä disnin di</i>

It is easy to notice that *gartəm* etc. reflect CA present forms, *kardam* etc. The stressed *a* turns to *ə*, e. g. *amaṛn* ‘summer’ > *amɔṛ*, *ardar* ‘righteous’ > *artɔr*, *beran* ‘mouth’ > *pərɔn*, *car* ‘tree’ > *jɔṛ*. Without the particle *gä*, the forms are used for the subjunctive, thus: Beylan *gartəm* ‘I may read’ from CA *kardam* ‘I read, I am reading’. The conditional and debititive are formed with *gə⁸* and *bə / bədə* respectively: *gə gartəm* ‘I will read’ and *bə / bədə gartəm* ‘I have to read’. Table 10 combines the aorist paradigm of the verb *tam* ‘to give’ in Middle Armenian (1901: 333, 2002: 315–316; Ant‘osyan 1975: 213; L. S. Hovsep‘yan 1997: 68–69), Modern West and East Armenian (Ant‘osyan 1975: 214), Svedia (Ačaryan 2003: 494, 498), Aramo (Łaribyan 1958: 47) and Zeyt‘un (Ačaryan 2003: 243).⁹

⁷ Compare Hamšen 2nd-person singular imperfect ending *-di*, e. g. *kie-di gu* ‘you were writing’ (on this see Vaux 2007: 268).

⁸ Łaribyan (1955: 229) explicitly distinguishes between *gä* and *gə*.

⁹ In the table (and henceforth) CA = Classical Armenian, SWA = Standard Western Armenian.

Table 10: Aorist paradigms in selected dialects

	CA	Middle SWA	Aramo	Proto.-Aramo	Svedia	Proto-Sved.	Zeyt'un	Proto-Zeyt'un
	Arm.							
SG	<i>etu</i>	<i>tu-i</i>	<i>tu-i</i>	<i>ədva</i>	* <i>etu-i</i>	<i>dva</i>	* <i>tu-i</i>	* <i>tu-i</i>
	<i>etur</i>	<i>tu-ir</i>	<i>tu-ir</i>	<i>ədvəy</i>	* <i>etu-e(r)</i>	<i>dvir</i>	* <i>tu-er</i>	* <i>tu-ir</i>
	<i>et</i>	<i>etu-r,</i> <i>eret,</i> <i>tu-aw</i>	<i>tu-aw</i>	<i>ida</i>	* <i>etu</i>	<i>i-dör</i>	* <i>e-tu-r</i>	* <i>tu-aw</i>
PL	<i>tuak'</i>	<i>tuak'</i>	<i>tu-ink'</i>	<i>ədvunk'</i>	* <i>etua(n)k'</i>	<i>dvunk'</i>	* <i>tu-ank'</i>	* <i>tu-ank'</i>
	<i>etuk'</i>	<i>tuik'</i>	<i>tu-ik'</i>	<i>ədväk'</i>	* <i>etu-ik'</i>	<i>dväk'</i>	* <i>tu-ik'</i>	* <i>tu-ik'</i>
	<i>etun</i>	<i>tuin</i>	<i>tu-in</i>	<i>ədväyn¹⁰</i>	* <i>etu-in</i>	<i>dven</i>	* <i>tu-in</i>	* <i>tu-in</i>

Syntax of basic clauses

- (1) *murd-ə* *k-asi*
 man-DEF.ART IND-say.PRS.3SG
 ‘The man says’
 (Ačaryan 2003: 546, line 8 from bottom)
- (2) *murd-ə* *k-ar̩ni* *z-at*
 man-DEF.ART.NOM.SG IND-take.PRS.3SG DEF.ART.ACC.-that
 p 'ara-n
 money-DEF.ART
 ‘The man is taking that money’
 (Ačaryan 2003: 546, line 4 from bottom)
- (3) *hälivür-ə* *k-ar̩ni* *z-at*
 old man-DEF.ART.NOM.SG IND-take.PRS.3SG DEF.ART.ACC-that
 murd-ə
 man-DEF.ART.ACC.SG
 ‘The old man is taking that man’
 (Ačaryan 2003: 548, line 11)
- (4) *Astuc* *kə-təsna* *z-leyr*
 god IND-see.PRS.3SG DEF.ART.ACC-mountain
 ‘The God sees the mountain’
 (Ačaryan 2003: 533)

¹⁰ Printed as *ənväyn* in Łaribyan (1958: 47), probably a misprint.

6.2. Hamſen / Homſec‘ma (the Hamſen or North-Western intergroup)

There are three basic subgroups of the Hamſen population (Vaux 2000–2001: 48; 2007: 257; Bläsing 2003: 7):

- Western Hemſinli, who live in the Turkish province of Rize, speak Turkish, and are Sunni Muslim. They preserve a huge amount of words and proper names from the Armenian dialect of Hamſen (Bläsing 1992, 1995). They are the descendants of Islamized Armenians from Hamſen.
- Eastern Hamſen/Hemſin(li)/Homſe(n)c‘ik‘, who live in the vicinity of Hopa in the province of Ardvin (also elsewhere in Turkey, Central Asia and Europe), speak Homſec‘ma (often being unaware that it is Armenian), and are Sunni Muslim.
- Northern Homſenc‘ik‘, who live in Abkhazia and southern Russia (Krasnodar), speak Homſec‘ma, and are Christian. They are the descendants of non-Islamized Hamſen Armenians who lived in the provinces of Samsun, Ordu, Giresun and Trabzon.

The most complete description of the dialect is Ačařyan (1947). Texts: Haykuni (1892) and Dumézil (1964, 1965, 1967, 1986–1987). Hamſen/Homſec‘ma is one of the most divergent Armenian dialects. Due to its isolated position the dialect preserves a considerable number of archaisms, both morphological and lexical, and also developed numerous peculiar innovations.

Phonology

The most characteristic phonological feature of Hamſen Armenian is the change *a* > *o* when followed by a nasal consonant: *amis* > *ɔmis* ‘month’, *ban* > *pɔn* ‘thing’, etc.; also in Turkish loans, e. g. *tavan* > *t'avɔn* ‘ceiling’ (Ačařyan 1947: 11, 22–24; Ačařyan 1971: 404–406; Bläsing 2003: 9–12). See also §6.1 (Svedia).

Nominal morphology

Table 11 gives the paradigms of *hac*‘, *i*-stem ‘bread’ and *mard*, *o*-stem ‘man, human’ in Hamſen (Ačařyan 1947: 90–91).

Table 11: Nominal paradigms in CA and Hamšen

	CA	Hamšen ¹¹		CA	Hamšen
SG	N Acc GD Abl I	<i>hac'</i> <i>(z)hac'</i> <i>hac'-i</i> <i>i hac'-ē</i> <i>hac'-i-w</i>	<i>hac'</i> <i>hac'</i> <i>hac'-i, -ec'</i> <i>hac'-ā, -ən</i> <i>hac'-ɔv</i>	PI Pl	<i>hac'-k'</i> <i>(z)hac'-s</i> <i>hac'-i-c'</i> <i>i hac'-i-c'</i> <i>hac'-i-w-k'</i>
					<i>hac'-er</i> <i>hac'-er</i> <i>hac'-er-u</i> <i>hac'-er-ā</i> <i>hac'-er-ɔv</i>
SG	N Acc GD Abl I	<i>mard</i> <i>(z)mard</i> <i>mard-oy</i> <i>i mard-oy</i> <i>mard-o-v</i>	<i>mašt</i> <i>mašt</i> <i>mašt-u</i> <i>mašt-ā(mä)</i> <i>mašt-ɔv</i>	PI Pl	<i>mard-k'</i> <i>(z)mard-s</i> <i>mard-oc'</i> <i>i mard-oc'</i> <i>mard-o-v-k'</i>
					<i>mašt-er</i> <i>mašt-er</i> <i>mašt-er-u</i> <i>mašt-er-ā(mä)</i> <i>mašt-er-ɔv</i>

A number of Classical Armenian words that were mainly used in the plural appear in the majority of dialects with petrification of the nominative plural form. Table 12 provides a few remarkable relics of the old genitive plural in Hamšen, Zeyt‘un¹² and Evdokia¹³ (see also §6.3. Artrial).

Table 12: Relics of old genitive plural in selected dialects

gloss	Case	CA	Hamšen	Evdokia	Zeyt‘un
eye	NOM.PL	<i>ač'-k'</i>	<i>ač'k</i>	<i>ač'k'</i>	<i>eč'vē</i>
	GEN.PL	<i>ač'-ac' & -ic'</i>	<i>ač'ic'(<-ic')</i>	<i>ač'ic'</i>	<i>eč'ɔc'(<-ac')</i>
wedding	NOM.PL	<i>harsan-i-k'</i>	<i>haysnik'</i>	<i>harsnik'</i>	<i>haysnək'</i>
	GEN.PL	<i>harsan-e-a-c'</i>	<i>haysnec'</i>	<i>harsnec'</i>	<i>haysnic'</i>
beard	NOM.PL	<i>mō/u/iru-k'</i>	<i>miruk'</i>	<i>mōruk'</i>	<i>muyuk'</i>
	GEN.PL	<i>mō/u/iru-a-c'</i>	<i>mirvac'</i>	<i>mōrvac'</i>	<i>muyvic'</i>

Another remarkable morphological relic is observable in the paradigm of the word for ‘star’, CA *astł*. The Hamšen dialect has nom.SG *astexk/k* ‘a star’, which is a frozen form of NOM.PL *astetk'*, and GEN.PL *astłec* ‘of stars’, cf. CA GEN.PL *astet-a-c'*. The GEN.SG *-ec'* is derived from the genitive plural ending *-eac'*, e. g. GEN.SG *gałnec'* from CA *kałneac'*, the GEN.PL form of *kałni* ‘oak’ (Ačařyan 1947: 91–94).

¹¹ Abl. *hac'-ā* is indefinite, and *hac'-ən* is definite. Other forms: *hac'-ec'-ən*, *hac'-ən*.

¹² See Ačařyan (1947: 98 and 2003: 188–189) and glossaries, also Martirosyan (2010: 17–18, 98–99, 454–455).

¹³ Gazančean (1899: 81) (the forms here do not seem phonetically very accurate).

Verbal morphology

Hamšen has the following important peculiarities in verbal morphology:

The CA infinitive ending *-(V)l* has been replaced by *-uš* (probably of Turkish origin) in all four conjugations, e. g. *berel* ‘to bring’ > *peruš*, *ert* ‘al ‘to go’ > *eštuš*, etc. (Ačaryan 1947: 11, 156–158; Ačaryan 1965: 46–47).

In Mala and the villages of Trapizon the verb ‘to have’ is used as a progressive marker: *b* ‘*erim uni* or *g-uni* ‘I am bringing’, *b* ‘*ereyə g-uni* ‘I was carrying’. Čanik uses a different pattern involving the infinitive and 3SG copula: *yes eguš ä* ‘I am coming’, *tun eguš ä* ‘you are coming’, etc.; or with a subject in genitive: *imə/ims eguš ä* ‘I am coming’, *k'ugə(d)* *eguš ä* ‘you are coming’, etc. (Ačaryan 1947: 11, 140–141).¹⁴

The second-person singular imperfect ending *-yd(ə) / -di* of Hamšen is comparable with *-idi* (of Turkish origin) seen throughout the whole imperfect paradigm in Hačən.¹⁵

First person plural imperfect: The Classical Armenian first person plural imperfect ending is characterized by the presence of a vowel *-a-*; thus, for example, in the *e*-conjugation we have the following set of endings: *-ēi* (1SG), *-ēir* (2SG), *-ēr* (3SG), *-ēak* (1PL), *-ēik'* (2PL), *-ēin* (3PL). In most dialects of Modern Armenian this *-a-* has been analogically eliminated, but some dialects have preserved it intact. In table 13 I present the Classical imperfect paradigm of the verb *utem* ‘to eat’ and its corresponding tense forms (imperfect or past subjunctive) in the dialects of Akn (Ačaryan 1911: 223, cf. 227 on Sebastia), Hamšen (Ačaryan 1947: 136, 139), and South Eastern Armenian. I also include two paradigms from the region of Dersim, *grem* ‘to write’ in Xarberd/Erznka, and *k'ašem* ‘to pull, drag’ in Č'mškacag (Bałramyan 1960: 22, 30); the latter preserves the *-a-* intact whereas the former is innovative.

Table 13: Imperfect paradigms in CA and in selected contemporary dialects

	CA	Hamšen	Akn	Č'mškacag	Xarb./Erz.	SEA
SG	<i>utēi</i>	<i>g-udeyə</i>	<i>g-üdei</i>	<i>gə k'ašeı</i>	<i>gə g'ərei</i>	<i>k-utei</i>
	<i>utēir</i>	<i>g-udeyd(ə)</i>	<i>g-üdeir</i>	<i>gə k'ašeir</i>	<i>gə g'əreir</i>	<i>k-uteir</i>
	<i>utēr</i>	<i>g-uder</i>	<i>g-üder</i>	<i>gə k'ašeı̄r</i>	<i>gə g'ərer</i>	<i>k-uter</i>
PL	<i>utēak'</i>	<i>g-udayk'ə</i>	<i>g-üdeank'</i>	<i>gə k'ašeank'</i>	<i>gə g'əreink'</i>	<i>k-uteink'</i>
	<i>utēik'</i>	<i>g-udéyə</i>	<i>g-üdeik'</i>	<i>gə k'ašeik'</i>	<i>gə g'əreik'</i>	<i>k-uteik'</i>
	<i>utēin</i>	<i>g-udéyə</i>	<i>g-üdein</i>	<i>gə k'ašein</i>	<i>gə g'ərein</i>	<i>k-utein</i>

The same 1PL *-a-* is seen in the aorist of Classical Armenian. Here again, some dialects preserve it intact, cf. e. g. the paradigms of *xatam* ‘to play’ and *sirem* ‘to love’

¹⁴ For a discussion, see Vaux (2007: 263).

¹⁵ See Ačaryan (1959: 568–569); cf. Vaux (2007: 268).

in Hamšen: *xalac ‘i* (1SG), *xalac ‘ir* (2SG), *xalac ‘* (3SG), *xalac ‘ak‘* (1PL), *xalac ‘ik‘* (2PL), *xalac ‘in* (3PL), *sirec ‘i* (1SG), *sirec ‘ir* (2SG), *sirec ‘* (3SG), *sirec ‘ak‘* (1PL), *sirec ‘ik‘* (2PL), *sirec ‘in* (3PL) (Ačaryan 1947: 128, 131). In old monosyllabic aorist forms of this dialect we find yet another archaism, the augment *e-*.¹⁶ Table 14 gives the example of *banam*, AOR stem *bac ‘-* ‘to open’ (Ačaryan 1947: 135).

Table 14: Aorist paradigm of ‘open’ in CA and Hamšen

	CA act.	CA med.	Hamšen
SG	<i>bac ‘i</i>	<i>bac ‘ay</i>	<i>pac ‘i</i>
	<i>bac ‘er</i>	<i>bac ‘ar</i>	<i>pac ‘ir</i>
	<i>e-bac ‘</i>	<i>bac ‘aw</i>	<i>pac ‘aw, ε-pac ‘</i>
PL	<i>bac ‘ak‘</i>	<i>bac ‘ak‘</i>	<i>pac ‘ak‘</i>
	<i>bac ‘ēk‘, -ik‘</i>	<i>bac ‘ayk‘</i>	<i>pac ‘ik‘</i>
	<i>bac ‘in</i>	<i>bac ‘an</i>	<i>pac ‘in</i>

A note on Hamšen *asuš*, AOR stem *ast-* ‘to say’

Classical Armenian *asem* ‘to say, tell, speak’ displays aorist forms based on *ast-* in Hamšen. According to Ačaryan (1947: 134–135), the *-t-* is an epenthesis of a phonetic nature, cf. *almas > almast* ‘diamond’, etc. This explanation is not entirely satisfactory because: (1) it is not clear why the *-t-* is only found in the aorist; and (2) we expect to see the reflexes of the aoristic *c-*. I, therefore, propose the following solution.

In Classical Armenian the aorist stem of this verb is *asac ‘*. This subtype displays both non-syncopated and syncopated aorist forms in Hamšen, namely *xalac ‘i* and *xalac ‘i* ‘I played’ (see Ačaryan 1947: 130–131). It seems, therefore, likely that Hamšen *asti* etc. derive from the syncopated forms **asc ‘i* etc.; the development **asc ‘i > asti* is probably due to dissimilation, which may have been triggered or reinforced by the *-(s)t-* epenthesis. For the inclusion of the 3SG.AOR.MED into the main paradigm, compare the aforementioned *pac ‘aw* alongside *ε-pac ‘*.

Table 15 combines the Classical and Hamšen paradigms with that of Šamaxi / K‘yärk‘yänj, an easternmost dialect (on which see Bałramyan 1964: 166).

¹⁶ For a discussion of this Hamšen archaism, see Vaux (2007: 265–266). It is also present in a neighboring dialect of Xotorjur (Kostandyan 1985: 60).

Table 15: The verb ‘to say’ in CA, Hamšen, and Šamaxi

	CA	Syncope	Hamšen	Šamaxi
SG	<i>asac‘i</i>	* <i>asc‘i</i>	<i>as-t-i</i>	<i>asc‘i, assi</i>
	<i>asac‘er</i>	* <i>asc‘er</i>	<i>as-t-ir</i>	<i>asc‘ir, assir</i>
	<i>asac‘</i>	* <i>as(a)c‘-aw</i> (med.)	<i>as-t-av</i>	<i>asec‘, asec</i>
PL	<i>asac‘ak‘</i>	* <i>asc‘ak‘</i>	<i>as-t-ak‘</i>	<i>asc‘ink‘, assink‘</i>
	<i>asac‘ēk‘, -ik‘</i>	* <i>asc‘ēk‘, -ik‘</i>	<i>as-t-ik‘</i>	<i>asc‘ik‘, assik‘</i>
	<i>asac‘in</i>	* <i>asc‘in</i>	<i>as-t-in</i>	<i>asc‘in, assin</i>

Table 16 combines the aorist paradigms of the irregular verb *tam* ‘to give’ in Hamšen (Ačaryan 1947: 137) and literary forms of ArmenianTable 16: The irregular verb *tam* ‘to give’ in Hamšen and literary forms of Armenian

	CA	Mid. Arm.	SWA	Hamšen
SG	<i>etu</i>	<i>tu-i</i>	<i>tu-i</i>	<i>dvi</i>
	<i>etur</i>	<i>tu-ir</i>	<i>tu-ir</i>	<i>dvir</i>
	<i>et</i>	<i>etu-r, eret, tu-aw</i>	<i>tu-aw</i>	<i>dvar, ered</i>
PL	<i>tuak‘</i>	<i>tuak‘</i>	<i>tu-ink‘</i>	<i>dvak‘</i>
	<i>etuk‘</i>	<i>tuik‘</i>	<i>tu-ik‘</i>	<i>dvik‘</i>
	<i>etun</i>	<i>tuin</i>	<i>tu-in</i>	<i>dvin</i>

Syntax of basic clauses

- (5) *hoxbey-s mezigi šad sirer gu*
uncle-my we.ACC much love.IMPF.3MS PARTICLE
‘My uncle loved us very much’
(Ačaryan 1947: 194, line 1)
- (6) *mek‘ unik‘ meg varjadun*
we.NOM have.1PL.PRS one school
‘We have got one school’
(Ačaryan 1947: 196, line 5)
- (7) *mir g‘elac‘i-k‘ gov šad behin gu*
we.GEN villager-NOM.PL cow.ACC.SG many keep.PRS.3PL PARTICLE
‘Our villagers keep many cows’
(Ačaryan 1947: 196 line 17–18)
- (8) *im ɔnun-s Ašot ä*
I.GEN name-my PN(MASC) COP.3SG
‘My name is Ašot’
(Ačaryan 1947: 199, line 3)

6.3. The Arťial (Transylvania) or extreme North-Western intergroup

Arťial (Transylvania)

This dialect had four branches: Poland, Hungary, Suč'ava and Romania. The dialect descriptions: Hanusz (1887–1889); Ačaryan (1953). In Pisowicz (2003), one finds a collection of Turkish, Arabic and Persian borrowings in Arťial.

Phonology

Voiced aspirates: for a discussion, see Pisowicz (1997: 216–217) and Vaux (1997: 233–234).

Nominal morphology

Important archaisms:

CA accusative definite marker *z-* (Ačaryan 1953: 10, 135–136, 156–157);

CA *harsaneac* ‘(GEN.PL of *harsani-k* ‘wedding’) > GEN *harsnic* ‘, LOC *harsnec* ‘*n*’ (see also §6.2. Hamšen);

CA *dustr*, GEN *dster* ‘daughter’ > *d*‘*usdrə*, GEN *d’sder* (Ačaryan 1953: 141–142).

Table 17 gives the paradigms of *hac* ‘, *i*-stem ‘bread’ and *mard*, *o*-stem ‘man, human’ in Arťial (Ačaryan 1953: 136–137).

Table 17: Nominal morphology in Arťial

		CA	Arťial			CA	Arťial
SG	N	<i>hac</i> ‘	<i>hac</i> ‘	PL	<i>hac</i> ‘- <i>k</i> ‘	<i>hac</i> ‘- <i>er</i>	
	ACC	(<i>z</i>) <i>hac</i> ‘	(<i>əz</i>) <i>hac</i> ‘		(<i>z</i>) <i>hac</i> ‘- <i>s</i>	(<i>əz</i>) <i>hac</i> ‘- <i>er</i>	
	GD	<i>hac</i> ‘- <i>i</i>	<i>hac</i> ‘- <i>i</i>		<i>hac</i> ‘- <i>i</i> - <i>c</i> ‘	<i>hac</i> ‘- <i>er</i> - <i>u</i>	
	ABL	<i>i</i> <i>hac</i> ‘- <i>ē</i>	<i>hac</i> ‘- <i>ē</i>		<i>i</i> <i>hac</i> ‘- <i>i</i> - <i>c</i> ‘	<i>hac</i> ‘- <i>er</i> - <i>e</i>	
	I	<i>hac</i> ‘- <i>i</i> - <i>w</i>	<i>hac</i> ‘- <i>ɔw</i>		<i>hac</i> ‘- <i>i</i> - <i>w</i> - <i>k</i> ‘	<i>hac</i> ‘- <i>er</i> - <i>ɔw</i>	

		CA	Arťial
SG	N	<i>mard</i>	<i>mard</i> ‘
	ACC	(<i>z</i>) <i>mard</i>	(<i>əz</i>) <i>mard</i> ‘
	GD	<i>mard</i> - <i>oy</i>	<i>mard</i> ‘- <i>u</i>
	ABL	<i>i</i>	<i>mard</i> ‘- <i>unme</i> ,
		<i>mard</i> - <i>oy</i>	<i>mard</i> ‘- <i>əven</i>
	I	<i>mard</i> - <i>o-v</i>	<i>mard</i> ‘- <i>ɔw</i> ,
			<i>mard</i> ‘- <i>əvɔw</i>

Verbal morphology

In the table below, I present the paradigms of *sirem* ‘to love’ in Ařtial (Ačařyan 1953: 160–164) and Polis (Ačařyan 1941: 149–151). The 2SG.IMPF ending *-r* is replaced by *-s* by analogy with the present. This did not take place in the Polish subdialect:

Table 18: Verbal morphology in Ařtial (‘love’)

Present			Imperfect		
	CA	Ařtial	Polis	CA	Ařtial
SG	<i>sirem</i>	<i>gi sirim</i>	<i>gə sirem</i>	<i>sirēi</i>	<i>gi sirei</i>
	<i>sires</i>	<i>gi siris</i>	<i>gə sires</i>	<i>sirēir</i>	<i>gi sireis</i>
	<i>sirē</i>	<i>gi sire</i>	<i>gə sire</i>	<i>sirēr</i>	<i>gi sirer</i>
PL	<i>siremk'</i>	<i>gi sirink'</i>	<i>gə sirenk'</i>	<i>sirēak'</i>	<i>gi sireink'</i>
	<i>sirēk'</i>	<i>gi sirik'</i>	<i>gə sirek'</i>	<i>sirēik'</i>	<i>gi sireik'</i>
	<i>siren</i>	<i>gi sirin</i>	<i>gə siren</i>	<i>sirēin</i>	<i>gi sirein</i>

The aorist (*sirec* ‘i ‘I loved’) has been replaced by the pattern [past participle + copula], *siril im*.

Syntax of basic clauses

- (9) *B'axd'-ə dun gu-g'a jəw*
husband -DEF.ART. home PARTICLE–come.PRS.3SG- and
gnganə hed ades gə zuruc 'e
wife.GEN.DAT.DEF.ART. with like.that PARTICLE talk.PRS.3SG
‘The husband is coming home and talks to (his) wife like this’ (*gu-g'a* < *gə*-u-particle + *ig'al* ‘to come’; *gnganə* < *gnig* ‘wife’; *gə zuruc'ε* < *gə*-particle + *zuruc'el* ‘to talk’)
(Ačařyan 1953: 223, line 9–10 from bottom)
- (10) *derder-ə inji asil ε*
priest-DEF.ART I.DAT tell.PRF.PTCP COP.3.SG
‘The priest has told me ...’
(Ačařyan 1953: 222, line 15 from bottom)
- (11) *xalp atves mə gə desnu mi ak'rav mə*
fraudulent fox a PARTICLE see.PRS.3SG one crow a
(< *gə*-particle + *desnul* ‘to see’)
‘A fraudulent fox is seeing a crow’
(Ačařyan 1953: 227, line 13)

6.4. Muš (the Muš / Tigranakert or South-Central group)

The Muš dialect was spoken in the territory of the Armenian historical province of Turuberan and in south of Ayrarat. During the Russo-Turkish wars, two large groups of the Muš and Alaškert populations migrated to Aparan and south of the Lake Sevan, now in the Republic of Armenia. Near Axalk‘alak‘, there are three villages of immigrants from Xnus. The 1915 Genocide survivors dispersed around the world.

Dialect descriptions: Patkanov (1875, 2), Mserianc (1897–1901), Ačařyan (1911: 116–139), Bałdasaryan-T‘ap‘alc‘yan (1958), Bałeš, village of Xult‘ik in Tarōnean (1961); Alaškert: Aparan/Aragac in Madat‘yan (1985).

Phonology

Fourfold system of plosives: *b*‘ (in anlaut only) *b p p‘*. CA *h-* becomes *x* in some subdialects, such as Bałeš and Xlat‘, and remains intact elsewhere.

In word internal and final positions, CA voiced aspirates become voiceless aspirates in Muš and voiceless unaspirated in Alaškert. In the same positions, CA voiceless aspirates become voiceless unaspirated in Alaškert but remain intact in Muš.

The dialect also has an initial voiced *h-*‘, either as prothetic or deriving from *y-* (cf. Martirosyan 2010: 764–765).

Nominal morphology

Archaisms: def. accusative *z-* and locative/allative *i /y-*; preposed prepositions: *məč ‘mər tan* ‘in our house’. The subdialect of Alaškert does not have the *z-*. Table 19 shows the paradigms of *p‘ayt*, *i*-stem ‘wood’ and *ji*, *o*-stem ‘horse’ in Alaškert (Madat‘yan 1985: 76–78).

Table 19: Nominal paradigms in the dialect of Alaškert

		CA	Alaškert	CA	Alaškert
SG	N	<i>p‘ayt</i>	<i>p‘ed</i>	PL	<i>p‘ayt-k‘</i>
	ACC	(<i>z</i>) <i>p‘ayt</i>	<i>p‘ed</i>		(<i>z</i>) <i>p‘ayt-s</i>
	GD	<i>p‘ayt-i</i>	<i>p‘ed-i</i>		<i>p‘ayt-i-c‘</i>
	ABL	<i>i p‘ayt-ē</i>	<i>p‘ed-en, -ic</i>		<i>i p‘ayt-i-c‘</i>
	I	<i>p‘ayt-i-w-</i>	<i>p‘ed-ɔv</i>		<i>p‘ayt-i-w-k‘</i>

		CA	Alaškert	CA	Alaškert
SG	N	<i>ji</i>	<i>j‘i</i>	PL	<i>ji-k‘</i>
	ACC	(<i>z</i>) <i>ji</i>	<i>j‘i</i>		(<i>z</i>) <i>ji-s</i>
	GD	<i>ji-oy</i>	<i>j‘i-u</i>		<i>ji-o-c‘</i>
	ABL	<i>i ji-oy</i>	<i>j‘i-u-c</i>		<i>i ji-o-c‘</i>
	I	<i>ji-o-v</i>	<i>j‘i-ɔv</i>		<i>ji-o-v-k‘</i>

Verbal morphology

Indicative present: *kə sirim* ‘I love’, *k-atlam* ‘I grind’. Table 20 combines the aorist paradigm of the verb *tam* ‘to give’ in Muš (Bałdasaryan-T‘ap‘alc‘yan 1958: 169), Sasun (Petoyan 1954: 59), Polis (Ačaryan 1941: 149) and elsewhere.

Table 20: Aorist paradigms for *tam* ‘to give’ in CA and South-Central dialects

	CA	SWA	SEA	Muš	Sasun	Polis
SG	<i>etu</i>	<i>tu-i</i>	<i>tv(ec')i</i>	<i>təv(ec')i</i>	<i>dväc'ə</i>	<i>duvi</i>
	<i>etur</i>	<i>tu-ir</i>	<i>tv(ec')ir</i>	<i>təv(ec')ir</i>	<i>dväc'ər</i>	<i>duvir</i>
	<i>et</i>	<i>tu-aw</i>	<i>tvec'</i>	<i>təvec'</i>	<i>dvəc'</i>	<i>duvav</i>
PL	<i>tuak'</i>	<i>tu-ink'</i>	<i>tv(ec')ink'</i>	<i>təv(ec')ink'</i>	<i>dväc'ək' (!)</i>	<i>duvink'</i>
	<i>etuk'</i>	<i>tu-ik'</i>	<i>tv(ec')ik'</i>	<i>təv(ec')ik'</i>	<i>dväc'ək'</i>	<i>duvik'</i>
	<i>etun</i>	<i>tu-in</i>	<i>tv(ec')in</i>	<i>təv(ec')in</i>	<i>dväc'ən</i>	<i>duvin</i>

Syntax of basic clauses

- (12) *Sarkis k-erta kə-hasni*
 Sargis (m. pers. name) PARTICLE-go.PRS.3SG PARTICLE-reach.PRS.3SG
Melikk'yand
 place name
 ‘Sargis is going and reaching Melikk'yand’ (*k-erta* < *kə*-particle + *ertal* ‘to go’; *kə-hasni* < *kə*-particle + *hasnel* ‘to reach’)
 (Madat‘yan 1985: 155, line 15 from bottom)
- (13) *Sarə-a-i kisur kə-tesna Sarkəs-in*
 Sarah-GEN husband’s mother IND-see.PRS.3SG Sargis-ACC
 ‘The mother-in-law (husband’s mother) of Sarə sees Sargis’ (*kə-tesna* < *kə*-particle + *tesnal* ‘to see’)
 (Madat‘yan 1985: 155 line 13 from bottom)
- (14) *Məla Nasradin šad b'amparag mar̄t er*
 Molla Nasreddin much lewd man be.IMPF.3SG
 ‘Molla Nasreddin was a very lewd man’
 (Madat‘yan 1985: 158 line 10)

6.5. Van (the Van or Southern intergroup)

Before the 1915 Genocide, the Van dialect and the Moks and Šatax subdialects, as well as the (sub)dialects of Ozim and Diadin were spoken in territories that historically belonged to Vaspurakan and Moks, provinces of Greater Armenia. The 1915 Genocide survivors dispersed around the world. Some populations settled in the Republic of Armenia.

Dialect description: Phonology of Van (Ačaryan 1903); Van – in the glossary Moks and Ozim are also included (Ačaryan 1911: 140–158; 1952); Šatax

(M. H. Muradyan 1962); Moks (M. H. Muradyan 1982); Sevan region (Bałramyan 1972, *passim*); Diadin/Vardenis (V. Xač‘atryan 2004). A very rich glossary of Moks: Orbeli (2002), ca. 4500 words. Ozim is described in Ačařyan (1911: 147–150), Łaribyan (1953: 93–97), N. E. Hovsep‘yan (1966), Arewikyan (1967), and especially in the unpublished thesis by Hovsep‘yan (1970, *non vidi*). Texts from Ozim: Kaycorik (1899); Ačařyan (1911: 156–158).

M. H. Muradyan (1982) discusses 72 phonological and morphological isoglosses between Urban Moks, Rural Moks, Ozim, Šatax and Van and concludes that Rural Moks is closest in relationship to Urban Moks and furthest from Van, and Šatax takes an intermediate position between Rural Moks and Van. According to some scholars (Łaribyan 1953: 93–95; N. E. Hovsep‘yan 1966), Ozim should be treated as a distinct dialect. Ĵahukyan (1972: 135) takes it together with Van.

Phonology

One of the earliest Armenian dialectal isoglosses (see §4.1) is Ačařyan’s Law, that is, the fronting of back vowels after voiced obstruents.¹⁷ This feature unites the Van-Urmia and Goris/Łarabał/Krzen/Agulis areas. Table 21 provides examples for both voiced and unvoiced obstruents in the initial and internal positions respectively.

Table 21: Fronting of back vowels after voiced obstruents (Ačařyan’s Law)

CA	gloss	Van	Moks/Šat.	Urmia	Łarabał	Krzen
bak	courtyard	päk	päk	päk ^y	päk, pæk	—
bah	spade	päx	päx	päx	päh	bäh
barak	subtle	päräk	pärak	päräk ^y	pä/erak	bärak
barj	pillow	pärc‘	pärc‘	pärc‘	pä/erc‘	bärj/c‘
berň	burden	per̄	pe/eř	per̄	peřnə	ber̄
beran	mouth	peran	peran	peran	péran	bəran
berd	fortress	per̄t‘	pirt‘/pert‘	per̄t‘	—	—
bołk	radish	pöök	pölk	poxk ^y	pö/exk	böök
boc‘	flame	pöc‘	püc‘	poc‘	—	—
burd	wool	pürt‘	pürt‘	pürt‘	pürt‘	bürd/t‘
pakas	less	pakas	pakas	pakas	pákas	pakas
pał	cold	pał	pał	pał	—	—
pařaw	old woman	pařav	pařav	pařav	pařav	pařav
port	navel	pořt	puřt	pořt	pɔr/řt	pɔrt
gal	to come	k'äl	k'äl	ik'äl	k'äl	g'öl
garň	lamb	k'är̄	k'är̄	k'är̄	k'ärnə	g'är̄
garun	spring	k'ärün	k'ärün	k'ärün	k'ärunk‘	g'ärunk‘
get	river	k'et	k'it/k'et	k'et	k'et	g'et

¹⁷ For a discussion and literature on Ačařyan’s Law, see Vaux (1992) and Martirosyan (2010: 747).

CA	gloss	Van	Moks/Šat.	Urmia	Łarabał	Křzen
<i>gort</i>	frog	<i>k'ört</i>	<i>k'ört</i>	<i>kört</i>	<i>k'ε/ɔ̄rt 'nuk</i>	<i>g'ört</i>
<i>kap</i>	tie	<i>kap</i>	<i>kap</i>	<i>kap</i>	<i>kap</i>	<i>kap</i>
<i>karōt</i>	need(y)	<i>karɔt</i>	<i>karɔt</i>	<i>karɔt</i>	<i>kárɔt</i>	<i>karɔt</i>
<i>ker</i>	food	<i>ker</i>	<i>kir/ker</i>	—	<i>ker-</i>	<i>ker</i>
<i>kot'</i>	handle	<i>kot'</i>	<i>kut /kot'</i>	<i>kot'</i>	<i>köt'</i>	<i>köt'</i>
<i>kut</i>	grain	<i>kut</i>	<i>kut</i>	<i>kut</i>	<i>köt</i>	<i>kut</i>
<i>datjn</i>	wild mint	<i>täxc'</i>	<i>täxc'</i>	<i>täxc</i>	<i>té(l)sna</i>	<i>dälj</i>
<i>das</i>	order	<i>tǟs</i>	<i>tǟs</i>	<i>tǟs</i>	<i>tǟs, tes</i>	<i>dǟs</i>
<i>detl</i>	remedy	<i>tel</i>	<i>tel/tel</i>	<i>tel</i>	<i>tel</i>	<i>detl</i>
<i>dotl</i>	tremor	<i>tȫl</i>	<i>tȫl</i>	<i>tȫl</i>	<i>tȫl</i>	<i>dȫl</i>
<i>tasn</i>	ten	<i>tas</i>	<i>tas</i>	<i>tassə</i>	<i>tásə</i>	<i>tasə</i>
<i>tel(i)</i>	place	<i>tel</i>	<i>tel/tel</i>	<i>tel</i>	<i>t(ə)etl</i>	<i>tel</i>
<i>tun</i>	house	<i>tun</i>	<i>tun</i>	<i>tun</i>	<i>tən</i>	<i>tu</i>

CA	gloss	Van	Moks	Šatax	Łarabał	Křzen
<i>ambar</i>	storehouse	<i>ämbär</i>	<i>ambär</i>	<i>ambär</i>	<i>ámbar</i>	<i>ambar</i>
<i>andundk'</i>	abyss	<i>andundk'</i>	<i>händütk'</i>	<i>händütk' y</i>	<i>əndðxtə</i>	—
<i>apranks'</i>	products	<i>apranks</i> y	<i>apranks</i> y	<i>apranks</i> y	<i>áprank'</i>	<i>apranks'</i>
<i>ardar</i>	righteous	<i>ärtär</i>	<i>artär</i>	<i>artär</i>	<i>ärt'är</i>	—
<i>aregakn</i>	sun	<i>arek'äk</i>	<i>Erek'äk</i>	<i>arek'äk</i>	<i>ərik'nak</i>	<i>äräg'äk</i>
<i>arcat'</i>	silver	<i>ařcat'</i>	<i>ar/řcat'</i>	<i>ařcat'</i>	<i>árcat'</i>	<i>arcat'</i>
<i>arjař</i>	cattle	<i>äčär</i>	<i>ačär</i>	<i>ačär</i>	<i>áč'a/eřnə</i>	<i>ajäř</i>
<i>arjasp</i>	vitriol	—	<i>arčäp's</i>	<i>arčäps</i>	—	—
<i>bambak</i>	cotton	<i>pämbäk</i>	<i>pämbäk</i>	<i>pämbäk</i>	<i>pém̄ba/äk</i>	<i>bän/mbäk</i>
<i>borbos</i>	mold/Schimmel	<i>peɔrpəcs</i>	<i>pürpürs</i>	<i>pürpös</i>	<i>p'erp'ěšn-</i>	<i>bämberň-</i>
<i>gagat'n</i>	summit	<i>käk'ät'</i>	<i>käk'ät'</i>	<i>käk'ät'</i>	—	<i>Šm. k'äk'ät</i>
<i>gangat</i>	complaint	<i>k'änk'ät</i>	<i>k'änk'ät</i>	<i>k'änk'ät</i>	<i>k'ěng'ät</i>	<i>k/g'äng'ät</i>
<i>dadarel</i>	to cease	<i>tätreł</i>	<i>tätäril</i>	<i>tätreł</i>	<i>t'ət'áre/il</i>	<i>t'ät'ärel</i>
<i>ddum</i>	pumpkin	<i>tə̄t'üm</i>	<i>tə̄t'üm</i>	<i>tə̄t'üm</i>	—	—
<i>erkank'</i>	hand-mill	<i>erkanck</i> y	<i>erkanck</i> y	<i>erkanck</i> y	<i>árkank'</i>	<i>arkank'</i>
<i>kanač'</i>	green	<i>kanač'</i>	<i>kanač'</i>	<i>kanač'</i>	<i>kána(n)č'</i>	<i>kananč'</i>
<i>kardal</i>	to shout, read	<i>kärtäl</i>	<i>kartäl</i>	<i>kartäl</i>	<i>kárt'e'il</i>	<i>kardal</i>
<i>karkut</i>	hail	<i>karkut</i>	<i>karkut</i>	<i>karkut</i>	<i>kárku/ɔt</i>	<i>karkut</i>
<i>margarit</i>	pearl	<i>mark'ərit</i>	<i>märk'ərit</i>	<i>märk'ərit</i>	—	<i>mərk'ärít</i>
<i>mardak</i>	beam (of wood)	<i>märtäk</i>	<i>martäk</i>	—	—	—
<i>šatgam</i>	turnip	<i>šäxk'äm</i>	<i>šaxk'äm</i>	<i>šäxk'äm</i>	<i>šaxkam</i>	—
<i>urbat'</i>	Friday	<i>urpät'</i>	<i>urpät'</i>	<i>urpät'</i>	<i>ʒrp'ät'</i>	<i>urbät'</i>
<i>paganel</i>	to kiss	<i>pak'el</i>	<i>pak'nil</i>	<i>pak'nel</i>	—	—

The vocalism in Moks and Šatax is worthy of note. In e. g. Van *ärtär* and Łarabał *ärt* ‘är ‘righteous’ we can reconstruct the following sequence of developments: CA *ardar* > **artär* > *ärtär*. Moks and Šatax¹⁸ did not share the assimilation in vowel quality with Van, but preserved the intermediate form *artär*.

Nominal morphology

There are a number of plural markers, such as (Ačaryan 1952: 108–115):

- -er, with monosyllables, e. g. *xac*‘, pl. *xac*‘er;
- -ner (1) with words that became monosyllabic after the loss of final -n, e. g. *cük* (from CA *jukn* ‘fish’), pl. *ckner*; (2) with polysyllables, e. g. *axper* (cf. CA *elbayr* ‘brother’), pl. *axper-ner*;
- -hyter < *-k-ter, with polysyllables ending in a vowel or -n, e. g. *k'ini* (< CA *gini* ‘wine’), pl. *k'ini-hyter*; *mařan* ‘pantry, cellar’, pl. *mařan-hyter*;
- -estar, e. g. *ik'i*, pl. *ek'estar* from CA *aygi* ‘vineyard’, cf. coll. *aygestan(i)*.

An archaism that is preserved in Moks, Šatax and Ozim but lost in Van is the def. accusative marker z-: *əz-car kətric*‘ ‘(he/she) cut the tree’; *əz-lač höröxic*‘ ‘(he/she) sent the boy’.

As Ačaryan (1952: 13) points out, the ablative in -ic‘ and the accusative/dative by-forms with animate nouns give Van an intermediate position between western and eastern dialects.

The ablative of place names is formed with -a, e. g. *Vänä* ‘from Van’, cf. CA abl. *i Vanay*.

Relics of Classical Armenian plural genitive, e. g. Van *cer*, gen. *cerac*‘ reflecting CA *jerñ* (cf. pl. *jer-k*) and gen.pl *jeř-a-c*‘ ‘hand’, respectively. Table 22 provides the paradigms of *hac*‘, *i*-stem ‘bread’ and *gini*, *wo*-stem, *ea*-stem ‘wine’ (Ačaryan 1952: 117, 121).

Table 22: Nominal paradigms in the Van or Southern intergroup

	CA	Van		CA	Van
SG	N	<i>hac</i> ‘	PL	<i>hac</i> ‘-k‘	<i>xac</i> ‘-er
	ACC	(<i>z-</i>) <i>hac</i> ‘		(<i>z</i>) <i>hac</i> ‘-s	<i>xac</i> ‘-er
	GD	<i>hac</i> ‘-i		<i>hac</i> ‘-i-c‘	<i>xac</i> ‘-er-ac‘
	ABL	<i>i hac</i> ‘-ē		<i>i hac</i> ‘-i-c‘	<i>xac</i> ‘-er-ac‘
	I	<i>hac</i> ‘-i-w		<i>hac</i> ‘-i-w-k‘	<i>xac</i> ‘-er-ɔv

¹⁸ See Ačaryan (1952: 248) and M. H. Muradyan (1962: 192). The form is corroborated by the genuine record of Orbeli (2002: 208) in the Moks area in 1911–1912.

		CA	Van	CA	Van
SG	N	<i>gini</i>	<i>k'ini</i>	<i>gine-a-c'</i>	<i>k'ini-hy-ter</i>
	ACC	<i>(z-)gini</i>	<i>k'ini</i>		<i>k'ini-hy-ter</i>
	GD	<i>ginw-oy</i>	<i>k'inu</i>		<i>k'ini-hy-ter-ac'</i>
	ABL	<i>i ginw-o-y</i>	<i>k'inu-c'</i>		<i>k'ini-hy-ter-ac'</i>
	I	<i>ginw-o-v,</i> <i>gine-a-w</i>	<i>k'in-ɔv</i>		<i>k'ini-hy-ter-ɔv</i>

Verbal morphology

Table 23 combines the paradigms of the verb *tam* ‘to give’ in Van (Ačařyan 1952: 173), Moks (M. H. Muradyan 1982: 173), Šatax (M. H. Muradyan 1962: 148), in CA and SWA.

Table 23: Verbal paradigms in the Van or Southern intergroup

	CA	SWA	Van	Moks	Šatax
SG	<i>etu</i>	<i>tu-i</i>	<i>tv(ic')i</i>	<i>təvə</i>	<i>təvi</i>
	<i>etur</i>	<i>tu-ir</i>	<i>tv(ic')ir</i>	<i>təvir</i>	<i>təvir</i>
	<i>et</i>	<i>tu-aw</i>	<i>tvec', i-tu(r)</i>	<i>i-tu</i>	<i>i-tu</i>
PL	<i>tuak'</i>	<i>tu-ink'</i>	<i>tv(ic')ink'</i>	<i>təvink'</i>	<i>təvink'</i>
	<i>etuk'</i>	<i>tu-ik'</i>	<i>tv(ic')ik'</i>	<i>təvik'</i>	<i>təvik'</i>
	<i>etun</i>	<i>tu-in</i>	<i>tv(ic')in</i>	<i>təvin</i>	<i>təvin</i>

Syntax of basic clauses

- (15) *im xɔr anun Märkär i*
 my father.GEN.SG name Margar COP.3SG.PRS
 ‘My father’s name is Margar’
 (Ačařyan 1952: 224, line 4)
- (16) *Xurik xat 'ün nstir er xɔv cari tak*
 Mrs. Xufik sit.PST.PTCP be.IMPF.3SG cool tree.GEN.DAT.SG under
 ‘Mrs. Xufik was sitting under a cool tree’
 (Ačařyan 1952: 214, line 14 from bottom)
- (17) *astl.er.n anus p 'elk' k-utan*
 star-NOM.PL.DEF sweet shine IND-give.PRS.3PL
 ‘The stars are shining sweetly’ (*k-utan* < *kə*-particle + *tal* ‘to give’)
 (Ačařyan 1952: 235, line 5)

B. Eastern grouping

6.6. Xoy-Urmia (the Xoy / Marała or South-Eastern group)

Jahukyan (1972: 135) here takes one dialect, Xoy/Marała, and identifies four representatives of it: Marała, Urmia, P'ayajuk and Xoy. This dialectal area comprises the territory of the Armenian historical province of *Parskahayk'* (Persian Armenia), as well as some adjacent regions, in particular Marała (east of Lake Urmia), where Armenians have lived since the 13th century. In 1828, a considerable part of the Armenian populations of these regions migrated to the Caucasus, the village of Aza (Naxijewan, on the River Arax), Łarabał and elsewhere. Before the First World War there lived 368 Armenian families in the Marała region, and in 1920 the number was 200. Nowadays the dialect is also spoken in a few regions of the Republic of Armenia, namely Artaşat, Vayoc'-jor, Kotayk', Hoktemberyan, etc. (Ačaryan 1911: 288; 1926: 9–13; 1951: 359; Łaribyan 1953: 342–343; Asatryan 1962: 9–11).

Dialect description of Urmia/Xoy in Ačaryan (1911: 288–290), Łaribyan (1939: 128–146; 1953: 342–358) and Asatryan (1962). Marała is described in Ačaryan (1926).

Phonology

Before a syllable-initial consonant, the velars *k/k'* and *g/g'* become *y*, while *k'/k''* becomes *hy*, cf. *taxtak* ‘board, plank’, pl. *taxtəyner*; *axč'ik'* ‘girl’, pl. *axč'iynēr*, *p'et'äg'* ‘beehive’, pl. *p'et'äyner*; *irek y* ‘three’, *irehy xet* ‘three times’, etc. This also holds for words with, e. g., *k'n* from *gn*, such as *aganim* ‘to put on clothes’ > *xahynel* (cf. *hak'ne'il* in the majority of dialects). See also §6.5, Van. On Ačaryan’s Law and the shift *h* > *x*, see §§5.1, 6.5 (Van), 6.8 (Łarabał).

Nominal morphology

There are three plural markers:

- -er in monosyllables, e. g. *ec* (from CA *ayc* ‘goat’), pl. *ec-er*;
- -ner (1) in words that became monosyllabic after the loss of final *-n*, e. g. *k'är* (< CA *garñ* ‘lamb’), pl. *k'ärner*; (2) in polysyllables, e. g. *k'äzän* (< CA *gazan* ‘beast’), pl. *k'äzänner*;
- -*k''er* (a combination of the Old Armenian plural markers *k'* and -*e(a)r*) in polysyllables, e. g. *pälis* (< CA acc.PL *banali-s* ‘key’), pl. *pälis-k''er*.

In some subdialects one finds the following distribution: -*ner* in words ending in a vowel, and -*k''er* in words ending in a consonant, cf. *ařu-ner* (*ařu* ‘brook’) vs. *karas-k''er* (*karas* ‘jar’).

In some subdialects the marker -*k''er* is used in possessive constructions for expressing the plurality of the possessor, cf. *tun-əs* ‘my house’, *tən-er-əs* ‘my houses’ and *tənə-k''er-əs* ‘our houses’. The paradigms of *car*, o-stem ‘tree’ and *hogi* ‘spirit’ in CA and the Xoy dialect (Asatryan 1962: 77–78) are provided in Table 24.

Table 24: Nominal paradigms CA and Xoy dialect

		CA	Xoy		CA	Xoy
SG	N	<i>car̄</i>	<i>car̄</i>	PL	<i>car̄-k‘</i>	<i>car̄-nēr</i>
	ACC	(z-)cař	cař		(z-)cař-s	cař-nēr
	GD	cař-o-y	cař-i(/ə)		cař-o-c‘	cař-nēr-i(/ə)
	ABL	i cař-o-y	cař-ē		i cař-o-c‘	cař-nēr-ē
	I	cař-o-v	cař-ɔv		cař-o-v-k‘	cař-nēr-ɔv

		CA	Xoy		CA	Xoy
SG	N	<i>hogi</i>	<i>xɔk‘i</i>	PL	<i>hogi-k‘</i>	<i>xɔk‘i-nēr</i>
	ACC	(z-)hogi	<i>xɔk‘i</i>		(z-)hogi-s	<i>xɔk‘i-nēr</i>
	GD	hogw-o-y	<i>xɔk‘-u</i>		hogw-o-c‘	<i>xɔk‘i-nēr-i(/ə)</i>
	ABL	i hogw-o-y	<i>xɔk‘iy-ε</i>		i hogw-o-c‘	<i>xɔk‘i-nēr-ε</i>
	I	hogw-o-v	<i>xɔk‘(iy)-ɔv</i>		hogw-o-v-k‘	<i>xɔk‘i-nēr-ɔv</i>

Verbal morphology

The dialect belongs to the *l*-group; the indicative present is formed as with the construction participle + copula: *üzel em* ‘I want’, *et‘äl em* ‘I go’ (Xoy); *üzel i* ‘I want’, *amč‘ənalá yim* ‘I am ashamed’ (Marała).

Remarkable is Urmia, in which the 3rd singular form has *-l* vs. *-s* in the rest of paradigm, thus: *kires em/s* ‘I/you write’, 3SG *kirel i* ‘he or she writes’; *mənas em/s* ‘I/you stay’, 3SG *mənal i* ‘he or she stays’. It is interesting to compare this pattern to that of Areš/Havarik (an eastern dialect displaying no consonant shift), in which the 1st person has *-li(s)*, whereas in the others the *-l-* is combined with the participial ending *-um*: 1SG *sirelis əm / sireli yəm*, 2SG *sirelam əs*, 3SG *sirelim i* (CA *sirem* ‘to love’).¹⁹ The paradigms of the present indicative of the verb *grem* ‘to write’ in Urmia and Areš/Havarik are provided in Table 25.

Table 25: Verbal paradigms CA and Urmia and Areš/Havarik dialects

	CA	Urmia	Areš/Hav.
SG	<i>grem</i>	<i>kires em</i>	<i>g'irelis əm /</i>
			<i>g'ireli yəm</i>
	<i>gres</i>	<i>kires es</i>	<i>g'irelam əs</i>
	<i>grē</i>	<i>kirel i</i>	<i>g'irelim i</i>
	<i>gremk‘</i>	<i>kires enk‘</i>	<i>g'irelam ank‘</i>
	<i>grēk‘</i>	<i>kires ek ə</i>	<i>g'irelam ak‘</i>
	<i>gren</i>	<i>kires en</i>	<i>g'irelam an</i>

¹⁹ For the forms and a discussion, see the relevant dialect descriptions, particularly Asatryan (1962: 7–17), as well as G. Geworgyan (1985, 2010).

The imperfect is formed by addition of the particle *er* to the present form, but the 3SG form is not distinct here. Thus:

- *k'ires em* ‘I write’ vs. *k'ires em er* ‘I was writing’;
- *k'ires es* ‘you write’ vs. *k'ires es er* ‘you were writing’;
- *k'irel i* ‘he/she writes’ vs. *k'ires er* ‘he/she was writing’.

The imperfect in Areš is composed of *-ləm* + copula: *g'ireləm ey* ‘I was writing’, *g'ireləm eyr* ‘you were writing’, *g'ireləm er* ‘he/she was writing’, etc.

The debititive is formed with the particle *ti* (< *piti*), e. g. *ti k'irem* ‘I have to write’. Remarkable is the by-form *tikyi / tikya*. A very interesting feature of this dialect is the use of a possessive suffix with verbs, e. g. *me ci p'rnenk'-t* (one horse let's.catch-2SG) ‘let's catch a horse for you’. This is due to Persian influence, cf. *dīdam-at* ‘I saw you’.

Syntax of basic clauses

- (18) *et växkuy Názär-ə me k'əšer kəngə xətə*
 that coward Nazar-DEF.ART one night wife with
semk'y-n ə tüs ik'älə
 threshold-DEF.ART COP.3SG out come.PRS.3SG
 ‘One night, this coward Nazar steps out onto the threshold’ (*kəngə* < *kənik'*
 ‘wife’; *ik'älə* < *ik'äl* ‘to come’)
 (Asatryan 1962: 170, line 8–9 from bottom)
- (19) *menk' y ləsir enk' y er k' yə zärzändeli anun-ə*
 we hear 1PL.PST.PRF your (SG) fearsome name-DEF.ART
 ‘We had heard your fearsome name (fame)’ (*ləsel* ‘to hear, listen to’)
 (Asatryan 1962: 175, line 5)
- (20) *axč'ik'-ə an-kərəl cəmerük' i*
 girl-DEF.ART un-cut watermelon COP.3SG
 ‘The girl is an uncut watermelon’
 (Asatryan 1962: 183, line 9)

6.7. Ararat (the Ararat or North-Eastern group)

This is one of the largest Armenian dialects. It is mainly spoken in the northern and central parts of the Republic of Armenia, historically the eastern part of the province of Ayrarat, as well as the territory of Taširk‘, a district in the province of Gugark‘.

Dialect description: Lori (Asatryan 1968), Hoktemberyan region (Bałdasarjan-T'ap'alc'yan 1973), Aštarak (Kapancjan 1975), general (Markosyan 1989). Dictionary of the Ararat dialect: Nawasardean (1903).

Phonology

Ararat is a typically eastern dialect with indicative present in *um* (*sirə/um em* ‘I love’) and penultimate accentuation. The series *g/k/k'* practically remains intact; some subdialects have voiced aspirates *b'/g'/d'* in initial position.

The diphthongs *oy* and *iw* both become *u* or (in e. g. Lori) *i*, e. g. *loys* ‘light’ > *lus* or *lis*, *k'oyr* ‘sister’ > *k'ur* or *k'ir*; *ariwn* ‘blood’ > *arun* or *arin*, *jiwn* ‘snow’ > *jun* or *jin*.

Nominal morphology

The paradigms of *hac*‘, *i*-stem ‘bread’, and *mard*, *o*-stem ‘man, human’ in the Ararat dialect (Ačaryan 1957: 549–551; Markosyan 1989: 109) are provided in Table 26.

Table 26: Nominal paradigms in the Ararat dialect

		CA	Ararat		CA	Ararat
SG	N	<i>hac</i> ‘	<i>hac</i> ‘	PL	<i>hac</i> ‘- <i>k</i> ‘	<i>hac</i> ‘- <i>er</i>
	ACC	(<i>z</i>) <i>hac</i> ‘	<i>hac</i> ‘		(<i>z</i>) <i>hac</i> ‘- <i>s</i>	<i>hac</i> ‘- <i>er</i>
	GD	<i>hac</i> ‘- <i>i</i>	<i>hac</i> ‘- <i>i</i>		<i>hac</i> ‘- <i>i</i> - <i>c</i> ‘	<i>hac</i> ‘- <i>er-i</i>
	ABL	<i>i</i> <i>hac</i> ‘- <i>ē</i>	<i>hac</i> ‘- <i>ic</i> ‘		<i>i</i> <i>hac</i> ‘- <i>i</i> - <i>c</i> ‘	<i>hac</i> ‘- <i>er-ic</i> ‘
	I	<i>hac</i> ‘- <i>i-w</i>	<i>hac</i> ‘- <i>ɔv</i>		<i>hac</i> ‘- <i>i-w-k</i> ‘	<i>hac</i> ‘- <i>er-ɔv</i>

		CA	Ararat
SG	N	<i>mard</i>	<i>mart</i> ‘
	ACC	(<i>z</i>) <i>mard</i>	<i>mart</i> ‘(- <i>u</i>)
	GD	<i>mard</i> - <i>oy</i>	<i>mart</i> ‘- <i>u</i>
	ABL	<i>i</i> <i>mard</i> - <i>oy</i>	<i>mart</i> ‘- <i>uc</i> ‘
	I	<i>mard</i> - <i>o-v</i>	<i>mart</i> ‘- <i>ɔv</i>

Verbal morphology

The debitive *pəti* (< *piti*) can be both preposed and postposed, in the latter case being conjugated, e. g. *pəti ałam* vs. *ałal (pə)tem* ‘I have to grind’.

One of the most interesting features of the Lori subdialect is the doublet forms of imperfect, one of which is identical with the CA paradigm and can be derived from it, though a secondary creation is also considered possible (Asatryan 1968: 24–25, 121–122; Jahukyan 1972: 105). Table 27 gives the relevant forms in CA and the Lori subdialect.

Table 27: Verbal paradigms for the imperfect in CA and Loři subdialect

Present		Imperfect			
	CA	Loři	CA	Loři 1	Loři 2
SG	<i>berem</i>	<i>berəm em</i>	<i>berēi</i>	<i>berei</i>	<i>berəm i</i>
	<i>beres</i>	<i>berəm es</i>	<i>berēir</i>	<i>bereir</i>	<i>berəm ir</i>
	<i>berē</i>	<i>berəm a</i>	<i>berēr</i>	<i>berer</i>	<i>berəm er</i>
PL	<i>beremk'</i>	<i>berəm enk'</i>	<i>berēak'</i>	<i>bereink'</i>	<i>berəm ink'</i>
	<i>berēk'</i>	<i>berəm ek'</i>	<i>berēk'</i>	<i>bereik'</i>	<i>berəm ik'</i>
	<i>beren</i>	<i>berəm en</i>	<i>berēin</i>	<i>berein</i>	<i>berəm in</i>

Similarly, there are doublet forms for the 3SG.AOR. ending in Loři, e. g. *ber-uc'* vs. *ε-ber* ‘he/she brought’. The latter is an archaic relic of CA *e-ber*. For the complete aorist paradigm, see §6.8, Łarabał.

Syntax of basic clauses

Łormacik Pōłosə (subdialect of Astapat)

- (21) *t̥ormacik Pōłos-ə k-aser*
deceased Pōłos-DEF.ART IND-say.IMPF.3SG
'The deceased Pōłos was saying'
(Ačaryan 1911: 48, line 9 from bottom)
- (22) *mənk' vec' k'ur ink' hama axper č'-unink'*
we six sister were but brother NEG-have.IMPF.1PL
'We were six sisters but did not have brothers'
(Ačaryan 1911: 49, line 7)
- (23) *t'ak 'avor-ə gənəm a bat-ə*
king-DEF.ART go-3SG.PRS garden-DEF.ART
'The king is going to the garden'
(HayžołHek' 8, 1977: 443, line 4 from bottom)

6.8. Łarabał (the Łarabał / Šamaxi or extreme North-Eastern intergroup)

This is one of the largest Armenian dialects. It mainly occupies the territory of the Armenian historical provinces of Arc'ax and Utik'. It is mainly spoken in Republic of Artsakh/Arc'ax (Nagorno-Karabakh Republic). The population here is now ca. 140,000. Some of the subdialects of Łarabał are spoken in surroundings, such as Goris/Syunik', south of the Republic of Armenia.

Dialect description: Łarabał proper (Ačaryan 1899; Davt'yan 1966); Goris/Syunik' (Vartapetjan 1962, Margaryan 1975); Tavuš (Mežunc' 1989); Šałax (Łaribyan 1939: 50–58; Davt'yan 1966); Hadrut' (Łaribyan 1939: 107–127; Połosyan 1965; Davt'yan 1966). Historical phonology of the dialects of the Syunik' and Arc'ax areas: (Ałabekyan 2010). Copious dictionary: (Sargsyan 2013).

Phonology

An archaic set of isoglosses such as accent retraction onto penultimate syllable and Ačaryan's Law (fronting of back vowels after voiced obstruents)²⁰ demonstrates an opposition between the South-Eastern periphery (Łarabał/Agulis area) and Van, on the one hand, and the Central and Western regions, on the other. The Van-Urmia intermediate area may originally have formed part of the South-Eastern area. Note also the sharp contrast between Muš and Łarabał (see section 6). Łarabał and Van also share the devoicing consonant shift, in this case in contrast with Agulis, Melri and Křzen.

The evidence for Ačaryan's Law (see the table in 6.5, Van, as well as Martirosyan 2010: 747) corroborates the assumption of A. E. Xač'atryan (1984: 321–322) that Łarabał *pérp'el* is the regular outcome of Classical Armenian *borb-* ‘to inflame’ rather than an archaic reflex of an otherwise unattested *e*-grade form **berb-*. Similarly untenable is the derivation of Łarabał *kétel* and Melri *gétil* ‘to hide, conceal oneself’ from an archaic **gel-*. A. E. Xač'atryan (1984: 321) convincingly argues that these forms rather continue *gałel* ‘to hide’ through Ačaryan's Law.

Nominal morphology

Morphological archaisms: Łarabał gen. *térnə* from CA *te-arñ*, gen. of *tēr* < **ti-ayr* ‘master, lord’ (Martirosyan 2010: 61); Łarabał *cerɔk'* from CA INSTR.PL *jer-a-w-k'* ‘by/with hands’ (Davt'yan 1966: 107). Table 28 shows the paradigm of *hac* ‘bread’ (Davt'yan 1966: 105, 114–118):

Table 28: Nominal paradigms in Łarabał

		CA	Łrb. indef.	Łrb. def.
SG	N	<i>hac'</i>	<i>hac'</i>	<i>hac'-ə</i>
	ACC	<i>(z-)hac'</i>	<i>hac'</i>	<i>hac'-ə</i>
GD	<i>hac'-i</i>	<i>hac'-u</i>	<i>hac'-(ə-)en</i>	
ABL	<i>i hac'-ē</i>	<i>hac'-a</i>	<i>hac'-a-n</i>	
I	<i>hac'-i-w</i>	<i>hac'-a-v</i>	<i>hac'-a-v-ə</i>	
LOC	<i>i hac'-i</i>	<i>hac'-um,</i> <i>hac'u mɛč'in</i>	<i>hac'-um-ə,</i> <i>hac'en mɛč'in</i>	

²⁰ For Ačaryan's Law in Łarabał see Ałabekyan (2010: 154).

		CA	Łrb. indef.	Łrb. def.
PL	N	<i>hac</i> ‘-k‘	<i>hac</i> ‘-er	<i>hac</i> ‘-er-ə
	ACC	(<i>z-</i>) <i>hac</i> ‘-s	<i>hac</i> ‘-er	<i>hac</i> ‘-er-ə
	GD	<i>hac</i> ‘-i-c‘	<i>hac</i> ‘-er-e/u	<i>hac</i> ‘-er-e-n
	ABL	<i>i hac</i> ‘-i-c‘	<i>hac</i> ‘-er-a	<i>hac</i> ‘-er-a-n
	I	<i>hac</i> ‘-i-w-k‘	<i>hac</i> ‘-er-a-v	<i>hac</i> ‘-er-a-v-ə
	LOC	<i>i hac</i> ‘-s	<i>hac</i> ‘-er-əm, <i>hac</i> ‘ere meč‘in	<i>hac</i> ‘-er-əm-ə, <i>hac</i> ‘eren meč‘in

Verbal morphology

Łarabał basically belongs to the *um*-branch: indicative present *sírəm əm* (also in Tavuš), Goris *sírum em* ‘I love’. Šałax and Hadrut‘, however, belong to (*l*) *is*-branch, cf. *mənayis əm* and *mənas əm* ‘I stay’, respectively.

CA paradigm of *ek*-, suppl. aor. of *gam* ‘to come’ (1SG *eki*, 2SG *ekir*, 3SG *ekn*, 3PL *ekin*, etc., imper. *ek*) is preserved in Łarabał aor. *yéke*, *yéker*, *yéka*, imper. *yek*, etc. (Martirosyan 2010: 249). The CA and Mid.Arm. aorist paradigms of the verb *tam* ‘to give’ are shown together with those in the dialects of Šamaxi (Bałramyan 1964: 129, 162, 167), Křzen (Bałramyan 1961: 151), Hadrut‘ (Połosyan 1965: 212) and Ararat/Lori (Asatryan 1968: 141) in Table 29.

Table 29: Verbal paradigms in CA and Mid. Arm., and selected dialects

	CA	Mid. Arm.	Šamaxi	Křzen	Hadrut‘	Lori
SG	<i>etu</i>	<i>tu-i</i>	<i>tur/vi</i>	<i>tə/uvı</i>	<i>tuve</i>	<i>təvi</i>
	<i>etur</i>	<i>tu-ir</i>	<i>tur/vir</i>	<i>tə/uvir</i>	<i>tuver</i>	<i>təvir</i>
	<i>et</i>	<i>etu-r, eret, tu-aw</i>	<i>u-tur, tur/vav</i>	<i>ə-ret, təvuc‘</i>	<i>tu/əvav</i>	<i>təvuc‘</i>
PL	<i>tuak‘</i>	<i>tuak‘</i>	<i>turink‘</i>	<i>tə/uvink‘</i>	<i>tuvek‘y</i>	<i>təvink‘</i>
	<i>etuk‘</i>	<i>tuik‘</i>	<i>turik‘</i>	<i>tə/uvik‘</i>	<i>tuvek‘y</i>	<i>təvik‘</i>
	<i>etun</i>	<i>tuin</i>	<i>turin</i>	<i>tə/uvin</i>	<i>tuven</i>	<i>təvin</i>

Syntax of basic clauses

- (24) *t'ak‘aver-ə həvak‘um a ir yerg‘ir-i imastun*
king-DEF.ART gather.PTCP-COP.3SG.PRS his country-GEN.SG wise
mart‘-unc‘-ə
man-ACC.PL-DEF.ART
'The king is gathering the wise men of his country'
(Davt‘yan 1966: 263, lines 11–12 from bottom)
- (25) *min t'ak‘aver mehet min araz a tehnum*
one king once one dream COP.3SG-see.PTCP
'Once upon a time a king sees a dream'
(Davt‘yan 1966: 265, line 16 from bottom)

- (26) *šinäc'i Sirk "visin Maral ənəmav min kov
villager Sargis-GEN.DAT Maral name-INS.SG one cow
unər
have.3SG.IMPF*
 ‘The villager Sargis had a cow named Maral’
 (Davityan 1966: 280, line 12)

Abbreviations

1,2,3	first, second, third person	MASC	Masculine
ACC	Accusative	NEG	Negation
COP	Copula	NOM	Nominative
DAT	Dative	PL	Plural
DEF	Definite	PRF.PTCP	Perfect participle
DEF.ART	Definite article	PRS	Present
GEN	Genitive	PST.PTCP	Past participle
IMPF	Imperfective	SG	Singular
IND	Indicative		

References

- Ačaryan, Hrač'ya Hakobi. 1898a. (Ačařean), K'nnut'iwn Aslanbēgi barbařin [Study of the dialect of Aslanbēg] *Bazmavēp* 56(1). 31–35; (2). 83–88.
- Ačaryan, Hrač'ya Hakobi. 1898b. (Ačařean), K'nnut'iwn Aslanbēgi barbařin [Study of the dialect of Aslanbēg] (reprinted from 1898). Venice: S. Lazar. (Usumnasirut'iwnner haykakan gawařabanut'ean, 1).
- Ačaryan, Hrač'ya Hakobi. 1899. (Ačařean), K'nnut'iwn Łarabali barbařin [Study of the dialect of Łarabał] (reprinted from “Ararat” 1899). Vařaršapat: Tparan Mayr At'oroy S. Ējmiačni. (Usumnasirut'iwnner haykakan barbařaxōsut'ean, 3).
- Ačaryan, Hrač'ya Hakobi. 1902. (Ačařean), T'urk'erēnē p'oxafeal bařer hayerēni mēj [Turkish loanwords in Armenian] (Ēminean azgagrakan žołovacu, vol. 3). Moscow: Vařaršapat.
- Ačaryan, Hrač'ya Hakobi. 1903. (Adjarian), Lautlehre des Van-Dialekts. *Zeitschrift für armenische Philologie* 1. 74–86, 121–138.
- Ačaryan, Hrač'ya Hakobi. 1909. (Adjarian), *Classification des dialectes arméniens*. Paris: Librairie Honoré Champion.
- Ačaryan, Hrač'ya Hakobi. 1911. (Ačařean), *Hay barbařagitut'iwn: uruagic ew dasaworut'iwn hay barbařneri (barbařagitan k'artēsov)* [Armenian dialectology: A sketch and classification of Armenian dialects (with a dialect map)]. Moskua: Nor-Naxijewan. (Ēminean azgagrakan žołovacu, vol. 8).
- Ačaryan, Hrač'ya Hakobi. 1913. (Ačařean), *Hayerēn gawařakan bařaran* [Armenian regional dictionary]. T'iflis (Ēminean azgagrakan žołovacu, vol. 9).
- Ačaryan, Hrač'ya Hakobi. 1925. (Ačařean), *K'nnut'iwn Nor-Naxijewani (Xrimi) barbaři* [Study of the dialect of Nor-Naxijewan (Crimea)]. Yerevan.

- Ačaryan, Hrač'ya Hakobi. 1926. (Ačafean), *K'nnut'iwn Marałayi barbari* [Study of the dialect of Marala]. Yerevan.
- Ačaryan, Hrač'ya Hakobi. 1935. (Ačafean), *K'nnut'iwn Agulisi barbari* [Study of the dialect of Agulis]. Yerevan: PetHrat (State Press).
- Ačaryan, Hrač'ya Hakobi. 1940a. (Ačafean), *K'nnut'iwn Nor-Jułayi barbari* [Study of the dialect of Nor-Jułaj]. Yerevan.
- Ačaryan, Hrač'ya Hakobi. 1940b-1951. (AčařHPatm 1-2): *Hayoc' lezvi patmut'yun* [History of the Armenian language], vol. 1 (1940), Yerevan: University Press; vol. 2 (1951). Yerevan: State Press of Armenia (HayPetHrat).
- Ačaryan, Hrač'ya Hakobi. 1941. *K'nnut'yun Polsahay barbari* [Study of the dialect of Polis Armenian]. *Gitakan ašxatut'yunner* of Yerevan State University, vol. 19. 19–250.
- Ačaryan, Hrač'ya Hakobi. 1947. *K'nnut'yun Hamšeni barbari* [Study of the dialect of Hamšen]. Yerevan: Academy Press.
- Ačaryan, Hrač'ya Hakobi. 1952. *K'nnut'yun Vani barbari* [Study of the dialect of Van]. Yerevan: University Press.
- Ačaryan, Hrač'ya Hakobi. 1953. *K'nnut'yun Ařtiali barbari* [Study of the dialect of Ařtial]. Yerevan: Academy Press (Hayeren barbarer: 1).
- Ačaryan, Hrač'ya Hakobi. 1955–1971. (AčařLiak), *Liakatar k'erakanut'yun hayoc' lezvi* [Complete grammar of the Armenian language] (Introductory vol. 1955); (vol. 3, 1957); (vol. 6, 1971). Yerevan: Academy Press.
- Ačaryan, Hrač'ya Hakobi. 1971–1979. (HAB), *Hayerēn armatakan bararan* [Armenian root dictionary], in 4 vols (2nd edn.). Yerevan: University Press. [The original edition: 1926–1935, in 7 volumes, Yerevan: University Press].
- Ačaryan, Hrač'ya Hakobi. 2003. *K'nnut'yun Kılıkiyai barbari* [Study of the dialect of Cilicia]. Yerevan: University Press.
- Ałababean, A. 1901. *Hawlabarc'oc' mtawor keank'ə* [Intellectual life of inhabitants of Havlabar]. T'iflis: Tparan K. Martiroseanc'i.
- Ałabekyan, Melada A. 2010. Syunik'-Arc'ax tarack'i barbarneri patmakan hnč'yunabnut'yun [Historical phonology of the dialects of the Syunik'-Arc'ax area]. *Hayereni barbaragitan atlas: usumnasirut'yunner ew nyut'er* 3, 54–165. Yerevan: "Nairi".
- Alaverdyan, S. A. 1968a. Ardvini barbari hnč'yunakan bnoroš aranjnahatkut'yunnerə [Phonological peculiarities of the dialect of Ardvin]. *Lraber hasarakakan gitut'yunneri* 1968, Nr 9. 83–94.
- Alaverdyan, S. A. 1968b. Ardvini barbari jewabanut'yan mi k'ani aranjnahatkut'yunnerə [Some morphological peculiarities of the dialect of Ardvin]. In: *PBH* 1968, Nr 3. 229–240.
- Ałayan, Ē. B. 1954. *Mełru barbar* [The dialect of Mełri]. Yerevan: Academy Press.
- Ałayan, Ē. B. 1958. Barbařayin hnaguyn tarberut'yunner hayerenum [Ancient dialectal differences in Armenian]. *Tetekagir: hasarakakan gitut'yunner* 1958, Nr 5. 67–80.
- Allahvērtean, Y. 1884. *Ulnia kam Zeyt'un*. K-Polis: Pałtatlean/Aramean.
- Alpōyačean, Aršak. 1952. (Archag Alboyadjian) *Patmut'iwn Ewdokioy hayoc': tełagrakan, patmakan ew azgagrakan telekut'iwnnerov*. Gahirē: "Nor astl" [History of the Armenians in Tokat. With topographical, historical and ethnographical information]. Cairo: New Star Press.
- Alpōyačean, Aršak. 1961. (Archag Alboyadjian or Alpoyačean) *Patmut'iwn Malat'ioy hayoc': tełagrakan, patmakan ew azgagrakan* [History of the Armenians in Malatia: with topographical, historical and ethnographical information.] Beirut: Sevan Press.

- Amatuni, Sahak Vardapet. 1912. *Hayoc ‘bař u ban* [Armenian word and speech]. Vałaršapat.
- Andranik. 1900. *Tērsim: čanaparhordut‘iwn ew tełagrut‘iwn* [Dersim: travel and topography]. T’iflis: Aragatip M. Martiroseanc‘i. (T’iflisi hayoc‘ hratarakč‘akan ənkerut‘iwn, 188).
- Andreasyan, Tigran. 1967. *Svediayi barbarə (K‘ist‘inəgi lezun)* [The dialect of Svedia]. Yerevan: Academy Press.
- Ant‘osyan, S. M. 1961. *Kesariayi barbarə* [The dialect of Kesaria]. Yerevan: Academy Press.
- Ant‘osyan, S. M. 1975. Xonarhman hamakargi zargac‘man ənt‘ac‘k‘ə. In *Hayoc‘ lezvi patmakan k‘erakanut‘yun*, vol. 2, 159–377. Yerevan: Academy Press.
- Arewikyan, a. i. 1967. Ozmi barbarə [The dialect of Ozim]. *Lraber hasarakakan gitut‘yunneri* 1967, Nr 5. 66–78.
- Asatryan, G. S. 1992. *Kontakty armjanskogo s novoiranskimi jazykami* [Contacts of Armenian with New Iranian languages]. (Doctoral dissertation). Yerevan: Academy Press.
- Asatryan, M. E. 1962. *Urmayıy (Xoyi) barbarə* [The dialect of Urmia (Xoy)]. Yerevan: University Press.
- Asatryan, M. E. 1968. *Lořu xosvack‘ə* [The subdialect of Loři]. Yerevan: “Mitk”, University Press.
- Avagyan, Serine Movsesi. 1978. Arčak. *Hay azgagrut‘yun ew banahusut‘yun* 8. Yerevan: Academy Press.
- Avagyan, Suren Artemi. 1973. *Vimakan arjanagrut‘yunneri hnč‘yunabanut‘yun* [Phonology of inscriptions] (X–XIV dd.). Yerevan: Academy Press.
- Awde, Nicholas (ed.). 1997. *Armenian perspectives: 10th anniversary conference of the Association Internationale des Etudes Arméniennes*. Richmond & Surrey: Curzon.
- Ayceamn. 1907. Nmoyš ma Aramō giwlin hay lezuēn [A sample of the Armenian language of the Aramo village]. *Handēs amsōreay* 21. 27.
- Avvazyan, Argam. 1984. *Agulis: patmamšakut‘ayin hušarjanner* [Agulis: historical-cultural memorials] Yerevan: “Hayastan”.
- Azatean, T‘. 1943. *Akn ew Aknc‘ik‘: azgagrakan u kensagrakan yišatakaran* [Akn: ethnographical and biographical monument]. Ist‘anpul: Tpagr. Y. Abrahamean.
- Bałdasaryan-T‘ap‘alc‘yan, S. H. 1958. *Mšo barbarə* [The dialect of Muš]. Yerevan: Academy Press.
- Bałdasaryan-T‘ap‘alc‘yan, S. H. 1973. *Araratyan barbarə xosvack‘nerə Hoktemberyani šrjanum* [The Ararat subdialects in the district of Hoktemberyan]. Yerevan: Academy Press.
- Bałramyan, R. H. 1960. *Dersimi barbarayin k‘artezə* [The dialectal map of Dersim]. Yerevan: Academy Press.
- Bałramyan, R. H. 1961. *Křeni barbarə* [The dialect of Křen]. Yerevan: Academy Press.
- Bałramyan, R. H. 1964. *Šamaxii barbarə* [The dialect of Šamaxi]. Yerevan: Academy Press.
- Bałramyan, R. H. 1972. *Hayereni barbarner: Sewana lči avazani xosvacqnerə* [Armenian dialects: the subdialects of the basin of Lake Sevan]. Yerevan: Haykakan petakan man-kavaržakan institut.
- Bałramyan, R. H. 1976. Xotrjuri barbarə [The dialect of Xotrjur] *PBH* 1976, Nr 3. 95–105.
- Barxudareanc‘, Makar Vard./Epis. 1883. Pələ-Pułi. T’iflis: Mov. Vardaneani ew Ənk. tparanum.
- Barxutareanc‘, Usta-Gēorg. 1883. *Arazə tarin ktari. Šuši*: Tparan M. Mahtesi-Yakobeanc‘.
- Baxtikean, S. A. 1934. *Arabkir ew šrjakayi giwlerə: patmakan-azgagrakan hamařōt tesut‘iwn* [Arabkir and the adjacent villages]. Pēyrut‘: Tp. “Vahagn” H. Yakobean.

- Beekes, Robert S. P. 2003. Historical phonology of Classical Armenian, apud *Kortlandt 2003*: 133–211.
- Bläsing, Uwe. 1992. *Armenisches Lehngut im Türkeitürkischen am Beispiel von Hemşin*. Amsterdam & Atlanta: Rodopi.
- Bläsing, Uwe. 1995. *Armenisch – Türkisch: Etymologische Betrachtungen ausgehend von Materialen aus dem Hemşingegebiet nebst einigen Anmerkungen zum Armenischen, insbesondere dem Hemşindialekt*. Amsterdam & Atlanta: Rodopi.
- Bläsing, Uwe. 2003. Some aspects of Armenian-Turkish language contacts in the Hamshen area. *Annual of Armenian linguistics* 22–23. 7–16.
- Bläsing, Uwe. 2007. Armenian in the vocabulary and culture of the Turkish Hemshinli. In Hovann Simonian (ed.), *The Hemshin: History, society and identity in the Highlands of Northeast Turkey*, 279–302. London & New York: Routledge.
- Čanikean, Y. K. 1895. *Hnut‘iwnk‘ Aknay* [Antiquities of Akn]. T‘iflis: Tparan M. D. Rötineanc‘i. (T‘iflisi hayoc‘ hratarakč‘akan ənkerut‘iwn, 89).
- Clackson, James. 2004–2005. Review of Kortlandt 2003 (with an appendix = Beekes 2003). *Annual of Armenian linguistics* 24–25. 153–158.
- Č‘olak‘ean, Yakob. 1986. *K‘esapi barbarə* [The dialect of K‘esap]. Haleb.
- Danielyan, T‘ewan. 1967. *Malat‘iayi barbarə* [The dialect of Malat‘ia]. Yerevan: Academy Press.
- Davt‘yan, K. S. 1966. *Leřnayin Łarabati barbarayin k‘arteza* [The dialectal map of the Mountainous Łarabat]. Yerevan: Academy Press.
- Dawit‘-Bēk, Melik‘ S. 1919. *Arabkiri gawařabarbařə: jaynabanakan ew k‘erakanakan usumnasiru iwn* [The dialect of Arabkir: a phonetic and grammatical study]. Vienna: Mxit‘arean Tparan. (Azgayin Matenadaran, 79).
- Dumézil, Georges. 1964. Notes sur le parler d’un Arménien musulman de Hemşin. *Académie Royale de Belgique: Mémoires de la Classe des Lettres* (Bruxelles) 57(4). 3–52.
- Dumézil, Georges. 1965. Notes sur le parler d’un Arménien musulman d’Ardala. *Revue des études arméniennes* (Paris) n. s. 2. 135–142.
- Dumézil, Georges. 1967. Trois récits dans le parler des arméniens musulmans de Hemşin. *Revue des études arméniennes* (Paris) n. s. 4. 19–35.
- Dumézil, Georges. 1986–87. Un roman policier en arménien d’Ardala. *Revue des études arméniennes* (Paris) n. s. 20. 5–27.
- Eremean, Aram. 1923. Č‘aharmahali hay žołovrdakan banahiwsut‘iwnə [Armenian folklore of Č‘aharmahal]. Vienna: Mxit‘arean tparan (Azgayin Matenadaran, 96).
- Eremean, Aram. 1930. *P‘šrank‘ner Jūlahay ew Hndkahay banahiwsut‘iwnic‘: tater ew žołovrdakan erger (17–19rd dar)* [Fragments of folklore of the Armenians of Julfa and India]. Vienna: Mxit‘arean Tparan (Azgayin Matenadaran, 129).
- Gabikean, Karapet. 1952. *Bargirk‘ Sebastahay gawařalezui* [Dictionary of the dialect of Sebastia]. Erusalēm: St Jacob.
- Gabriēlean, M. S. 1912. *Aknay gawařabarbařə* [The dialect of Akn]. Vienna: Mxit‘arean tparan.
- Galustean, Gr. H. (Kilikec‘i). 1934. *Maraš kam Germanik ew heros Zēyt‘un*. New York: “Koč‘nak” (Maraši hayrenakc‘akan miut‘iwn).
- Gasparyan, H. H. 1966. *Hajəni barbarə* [The dialect of Hačən]. Yerevan: Academy Press.
- Gasparyan, H. H. 1979. Dersim: patmaazgagrakan aknark [Dersim: a historical-ethno-graphical survey]. *PBH* 1979, Nr 2. 195–210.

- Gawařac‘i, H. M. H. 1903. *Erger, arakner, hanelukner, t‘erahawatut ‘iwnner... Xotrjroy* [Songs, fables, riddles, superstitions... of Xotorjür]. Tp‘xis.
- Gazančean, Yovhannēs. 1899. *Ewdokioy hayoc‘ gawařabarbařə* [The Armenian dialect of Tokat]. Vienna: Mxit‘arean tparan.
- Geworgyan, Gayane. 1985. *Mijčuļayin barbařnerə hayerenum* [Interdialects in Armenian]. *Lraber hasarakakan gitut‘yunneri* 1985, Nr 9. 52–60.
- Geworgyan, Gayane. 1994. (Gevorkian), The continuous present (շարունակական ներկայական) in Armenian dialects. *Annual of Armenian linguistics* 15. 43–62.
- Geworgyan, Gayane. 2010. Sahmanakan ełanaki nerka žamanaki taracakan-gorcařakan bnut‘agirə hayereni barbařnerum [Areal-functional characteristics of the indicative present tense in Armenian dialects]. In *Hayereni barbařagitakan atlas: usumna-sirut‘yunner ew nyut‘er* 3, 232–295. Yerevan: “Nairi”.
- Geworgyan, Geworg G. 1980. Łzlar. *Hay azgagrut‘yun ew banahyusut‘yun* 10. Yerevan: Academy Press.
- Geworgyan, T‘amar H. 1999. Lofi (Taširk‘ – Joraget). *Hay azgagrut‘yun ew banahyusut‘yun* 20. Yerevan: Academy Press “Gitut‘yun”.
- Greppin, John A. C. & Amalya A. Khachaturian 1986. *A handbook of Armenian dialectology*. Delmar & New York: Caravan Books.
- Grigorean (Łazarukean) Alēk‘ s & Sedrak‘ Garakēōzean. 1963. Yišatakaran Ėvērēk-Fēnēsēi. P‘ariz.
- Grigoryan, A. V. 1957. *Hay barbařagitut‘yan dasənt‘ac* [A handbook of Armenian dialectology]. Yerevan: University Press.
- Grigoryan-Spandaryan, M. M. 1971. *Leřnayin Łarabali banahyusut‘yunə* [Folklore of the Mountainous Łarabał]. Yerevan: Academy Press.
- Gurunyan, Hakob. 1991. *Hamšenahay banahyusut‘yun* [Folklore of the Hamšen Armenians]. Yerevan: “Luyս”.
- Guyumčean, Sanduxt. 1990. *Vařemi Marašə ew hayerə* [The ancient Maraš and the Armenians]. Los Angeles.
- Gyozalyan (Gełuni), Grigor Mat‘ewosi. 2001. *Musa leřan azgagrut‘yunə* [Ethnography of Musa-Ler] (Veržine Svazlyan, ed.). Yerevan: Academey Press “Gitut‘yun”.
- Hačean, H. M. 1907. *Hin avandakan hēk‘eat‘ner Xotorjroy* [Old traditional folk tales of Xotorjur]. Vienna: I vans Paštpan S. Astuacacni.
- Hakobyan, G. A. 1974. *Nerk‘in Baseni azgagrut‘yunə ew banahyusut‘yunə* [Ethnography and folklore of Lower Basen]. Yerevan: “Hayastan”.
- Halajyan, Geworg. 1973. Dersimi hayeri azgagrut‘yuno [Ethnography of the Armenians of Dersim]. *Hay azgagrut‘yun ew banahyusut‘yun* 5. Yerevan: Academy Press.
- HamšHamš. 2007. *Hamšēn ew Hamšenahayut‘iwn (gitažołovi niwt‘er)*. Yerevan, Pēyrut‘: Hamazgayin Vahē Sēt‘ean tparan.
- Hananyan, Galt‘uhı Karapeti. 1995. *Svediayi barbařə (Xrbeki xosvack‘)* [The dialect of Svedia (Subdialect of Xrbek)]. Yerevan: University Press.
- Haneyyan, A. N. 1978. *Tigranakerti barbařə* [The dialect of Tigranakert]. Yerevan: Academy Press.
- Haneyyan, A. N. 1982. Edesiyai barbařə [The dialect of Edesia]. *Hayereni barbařagitakan atlas: usumnasirut‘yunner ew nyut‘er* 1. 184–240.
- Haneyyan, A. N. 2010. Jayneł šnč‘ełnera hayereni barbařnerum [The voiced aspirates in the Armenian dialects]. (Part 1). In *Hayereni barbařagitakan atlas: usumnasirut‘yunner ew nyut‘er* 3, 166–231. Yerevan: “Nairi”.

- Hanusz, Johann. 1887–1889. Beiträge zur armenischen Dialectologie. *Wiener Zeitschrift für die Kunde des Morgenlandes* (WZKM) 1, 1887: 181–197, 281–313; WZKM 2, 1888: 63–70, 124–132, 291–308; WZKM 3, 1889: 38–50.
- Harut‘yunyan, Lewon. 1991. *Nšxarner Arc‘axi banahyusut‘yan* [Fragments of the Arc‘ax folklore]. Yerevan: University Press.
- Hayk, Vahē. 1959. *Xarberd ew anor oskečen daštə: yušamatean azgayin, patmakan, mšakut‘ayin ew azgagrakan* [Xarberd and its golden field]. New York.
- Haykuni, S. 1892. Trapizōni Zefanoz, Krōmila hay giwlōrēic‘ sovorut‘iwnner u bařbařa [The customs and dialect of the Armenian villages of Trapizon: Zefanoz and Krōmila]. *Ararat* 25(1). 428–447.
- HayLezBrbBař 1–7. 2001–2012. Hayoc‘ lezvi barbařayin bařaran [Dialectal dictionary of Armenian] (ed. A. Sargsyan et al.), vol. 1, 2001 [letters a-z]; vol. 2, 2002 [letters ē-c]; vol. 3, 2004 [letters k-č]; vol. 4, 2007 [letters m-č’]; vol. 5, 2008 [letters p-s]; vol. 6, 2010 [letters v-p’]; vol. 7, 2012 [letters p’-f]. Yerevan: Academy Press “Gitut‘yun”.
- HayžołHek‘ 1. 1959. *Hay žołovrdakan hek‘iat‘ner* [Armenian folk tales]. Yerevan: Academy Press.
- Hewsen, Robert H. 2001. *Armenia: a historical atlas*. Chicago and London: The University of Chicago Press.
- Hovannessian, Richard G. (ed.). 2000. *Armenian Van/Vaspurakan*. Costa Mesa, California: Mazda Publishers. (UCLA Armenian history and culture series, Historic Armenian cities and provinces, 1).
- Hovannessian, Richard G. (ed.). 2001. *Armenian Baghesh/Bitlis and Taron/Mush*. Costa Mesa, California: Mazda Publishers. (UCLA Armenian history and culture series, Historic Armenian cities and provinces, 2).
- Hovannessian, Richard G. (ed.). 2009. *Armenian Pontus: The Trebizond-Black Sea communities*. Costa Mesa, California: Mazda Publishers. (UCLA Armenian history and culture series, ‘Historic Armenian cities and provinces, 8).
- Hovsep‘yan, Hovhannes. 2009. *Łaradati hayerə* [The Armenians of Łaradał. Vol. 1: Azgařgrut‘yun (Suren Hobosyan, ed.); vol. 2: Banahyusut‘yun (Svetlana Vardanyan & Ėst‘er Xemč‘yan, eds.)]. Yerevan: Academy Press. (Azgagrakan ew banahyusakan žařangut‘yun, 1–2).
- Hovsep‘yan, Liana S. 1997. *ŽG dari hayeren jeřagreri hišatakaranneri lezun* [The language of the 13th-century colophons of Armenian manuscripts]. Yerevan. Academy Press & Matenadaran. (Accessing Armenian colophons, 2).
- Hovsep‘yan, Liana & Gayane Geworgyan. 2013. Armjanskie dialekty: obščij obzor [The Armenian dialects: General overview]. In Yuri B. Koryakov & Andrej A. Kibrik (eds.), *Languages of the World: Relict Indo-European languages of Western and Central Asia*, 320–334. Moscow: Academia.
- Hovsep‘yan, N. E. 1966. Ozmi barbaři mi k‘ani ařanjnahatkut‘yunneri masin [On a few peculiarities of the dialect of Ozim]. *PBH* 1966, Nr 2. 226–238.
- Žahukyan, G. B. 1972. *Hay barbařagitut‘yan neracut‘yun: vičakagrakan barbaragitut‘yun* [Introduction to Armenian dialectology: statistic dialectology]. Yerevan: Academy Press.
- Žahukyan, G. B. 1985 (Djahukian), Die archaischen Schichten armenischer Dialektwörter. In U. Pieper & G. Stickel (eds.), *Studia linguistica, diachronica et synchronica Werner Winter sexagenario anno MCMLXXXIII*, 151–160. Berlin & New York: Mouton de Gruyter.

- Čahukyan, G. B. 1987. *Hayoc' lezvi patmut'yun: naxagrayin žamanakašrjan* [History of the Armenian language: preliterary period]. Yerevan: Academy Press.
- Čalašyan, G. S. 2012. *Nor Naxijewani barbari bac 'atrakan bačaran* [Defining dictionary of the Nor Naxijewan dialect]. Rostov-na-Donu: "Starye russkie".
- JaynHamš 1–4. 1971–1999. *Jayn hamšenakan* [Hamšen voice]. Vol. 1 (Sero Xanzadyan, eds., Yerevan: "Hayastan", 1971); vol. 2 (A. Zeyt'yan, eds., Yerevan: "Sovetakan groł", 1979); vol. 3 (ed. by Sergey Vardanyan, Yerevan: "Xorhrdayin groł", 1989); vol. 4 (Sergey Galstyan & Lyudvig Karapetyan, eds., Yerevan: "Vernatun", 1999).
- K'alanč'ar, A. 1895. *Sasun*. [T'iflis].
- Kapančjan, Grigor. 1975. *Istoriko-lingvisticheskie raboty* [Historical-linguistic studies]. (Vol. 2). Yerevan: Academy Press.
- Kapri, Freiherr M. 1913. *Zwei Vorträge über die historische und kulturelle Bedeutung des armenischen Volkes: mit einem Vorwort des Wiener armenophilen Komitees und einem Anhang des Patriarchats-Memorandums, der Bevölkerungsstatistik und einer geographischen Karte von Armenien*. Wien: Mechtharisten.
- Karapetyan, Elijazar Karapeti. 1962. *Sasun: azgagrakan nyut'er* [Sasun: ethnographic materials]. Yerevan: State Press of Armenia (HayPetHrat).
- Karayan, Sarkis Y. 2000. Demography of Van province, 1844–1914. In Richard G. Hovannisian (ed.), *Armenian Van/Vaspurakan*, 195–208. Costa Mesa, California: Mazda Publishers.
- Karst, Josef (Joseph). 1901. *Historische Grammatik des Kilikisch-Armenischen*. Strassburg: Verlag von Karl J. Trübner ([1970] Photomechanischer Nachdruck. Berlin: Walter de Gruyter & Co.)
- Karst, Josef (Joseph). 2002. *Kilikyan hayereni patmakan k'erakanut'yun* [Historical grammar of Cilician Armenian]. (ModArm. translation from Karst 1901 by Dmitri Alivnyak). Yerevan: University Press.
- Kaycorik. 1899. P'ayal-p'ašan/alan u P'arušan. *Biwrakn* 17. 20–21, 119–120, 298.
- K'ēč'eān, Arak'ēl & Mkrtič Parsamean. 1952. *Akn ew Aknc'ik'* (initiated and compiled by Arak'ēl K'ēč'eān; composed and edited by Mkrtič Parsamean). Paris: Compatriotic Association of Eguen.
- Khachaturian, Amalya. 1983. The nature of voiced aspirated stops and affricates in Armenian dialects. *Annual of Armenian linguistics* 4. 57–62.
- Kortlandt, Frederik H. H. 1980. On the relative chronology of Armenian sound changes. In John Greppin (ed.), *Proceedings of the First international symposium on Armenian linguistics* (11–14 July, 1979), 97–106. Delmar, New York: Caravan Books.
- Kortlandt, Frederik H. H. 2003. *Armeniaca: comparative notes – with an appendix on the historical phonology of Classical Armenian* by Robert S. P. Beekes. Ann Arbor: Caravan Books.
- Kostandyan, D. M. 1979. *Erznkayi barbarə* [The dialect of Erznka]. Yerevan: Academy Press.
- Kostandyan, D. M. 1982. Xarberdi barbari hnč'yunabananakan ew jewabanakan himnakan afanjnahatkut'yunnerə [The main phonological and morphological peculiarities of the dialect of Xarberd]. In *Hayereni barbarəagitakan atlas: usumnasirut'yunner ew nyut'er* 1, 241–272. Yerevan: Academy Press.
- Kostandyan, D. M. 1985. Xotřjuri barbarayin yurahatkut'yunneri lezvašxarhagrakan bnut'agirə [The linguistic-geographical characteristics of the peculiarities of the Xotřjur dialect]. In *Hayereni barbarəagitakan atlas: usumnasirut'yunner ew nyut'er* 2, 46–68. Yerevan: Academy Press.

- Kuzub, Mixayil. 2007. Hamšēnc‘i k‘ristoneaneri hek‘iat‘ayin banahiwsut‘ean diwabanakan patkerac‘umnerə [The demonological notions of folklore of the Christians of Hamšen]. *HamšHamš*: 168–180.
- Łafadaryan, K. G. 1967. Nyut‘er Axalc‘xayi hay hamaynk‘i patmut‘yan veraberyal [Materials on the history of the Armenian community of Axalc‘xa]. *PBH* 1967, Nr 1. 117–124.
- Lalayean, E. 1892. (Lalayeanc‘), *Žawaxk‘i burmunk‘* [Scent of Žawaxk‘]. T‘iflis: Tparan “Arōr” T. Nazarean. (Žołovrdagrakan niwt‘er, 2).
- Lalayean, E. 1897. *Žawaxk‘* (reprinted from “Azgagrakan handēs”). T‘iflis: Tparan M. D. Rötineanc‘i.
- Lalayean, E. 1897a. Varanda: niwt‘er apagay usumnasirut‘ean hamar [Varanda: materials for future studies]. *Azgagrakan handēs* 2. 4–244.
- Lalayean, E. 1903. Borč‘alui gawař [District of Borč‘alu]. *Azgagrakan handēs* 10. 113–268.
- Lalayean, E. 1910. Vaspurakan: azgagrut‘iwn. *Azgagrakan handēs*, year 12, book 20. 116–196.
- Lalayean, E. 1912–14. (LalVasp 1–2), *Vaspurakan: banahyusut‘iwn* (reprinted from AzgHand). Vol. 1, 1912, T‘iflis: N. Ałaneanc‘; vol. 2, Šataxi barbařov, 1914, T‘iflis: “Esperanto”.
- Lalayean, E. 1914. Vaspurakan: hawatk‘. *Azgagrakan handēs*, year 15, book 25. 21–60.
- Lalayean, E. 1915. (Lalajanc), Vanskij vilajet: Vaspurakan. *SMOMPK* 44(2). 1–84.
- Lalayean, E. 1917a. Muš-Tarōn: azgagrut‘iwn. *Azgagrakan handēs*, year 16, book 26. T‘iflis [1916]. 149–194.
- Lalayean, E. 1917b. Vaspurakan: hawatk‘. *Azgagrakan handēs*, year 16, book 26, T‘iflis [1916]. 195–210.
- Lalayean, E. 1983–2004. *Erker*. Vol. 1, 1983 (Žavaxk‘); vol. 2, 1988 (Arc‘ax); vol. 3, 2004 (Borč‘alui gavař). Yerevan: Academy Press.
- Łaribyan, Ararat Sahaki. 1939. Hayeren barbařneri mi nor čyuł [A new branch of Armenian dialects]. *Gitakan ašxatut‘yunner Yerewani petakan hamalsarani* 11. 23–187.
- Łaribyan, Ararat Sahaki. 1953. *Hay barbařagitut‘yun: hnč‘yunabanut‘yun ew jewabannut‘yun* [Armenian dialectology: phonology and morphology]. Yerevan: Armenian State External [*heraka*] Pedagogical Institute Press.
- Łaribyan, Ararat Sahaki. 1955. Hayereni mi k‘ani norahayt barbařner [A few newly-found dialects of Armenian]. In *Gitakan ašxatut‘yunneri žołovacu* (Erevani X. Abovyan anvan Haykakan petakan mankavaržakan institut) 4, 195–241. Yerevan: Academy Press.
- Łaribyan, Ararat Sahaki. 1958. *Hayereni norahayt barbařneri mi nor xumb* [A new group of Armenian dialects]. Yerevan: Academy Press.
- Łaziyan, Alvard S. 1983. Arc‘ax. *Hay azgagrut‘yun ew banahyusut‘yun* 15. Yerevan: Academy Press.
- Łaziyan, Alvard & Svetlana Vardanyan. 2004. *Bak‘vi hayoc‘ banahyusut‘yunə*. Yerevan: “Zangak”.
- Lisic‘yan, Step‘an D. 1969. *Zangezuri hayerə* [The Armenians of Zangezur]. Yerevan: Academy Press.
- Lisic‘yan, Step‘an D. 1981. Leñayin Łarabalı hayerə: azgagrakan aknark [The Armenians of the Mountainous Łarabal: ethnographical survey]. *Hay azgagrut‘yun ew banahyusut‘yun* 12. 7–84.
- Lusenc‘, Ašot Łukasi. 1982. *Areši barbařə* [The dialect of Areš]. Yerevan: Academy Press.
- Madat‘yan, K‘. M. 1985. *Alaškerti xosvack‘ə (Aparan-Aragaci tarack‘)* [The subdialect of Alaškert]. Yerevan: Academy Press.

- Mak‘sudeanc‘, Mesrop. 1910. Aknay barbari p‘ox arac baferə ötar lezuneric‘ [The loanwords of the dialect of Akn]. *Handēs amsōreay*. 57–63.
- Mak‘sudeanc‘, Mesrop. 1911. Zeyt‘uni barbari jaynakān drut‘iwnə [The phonology of the dialect of Zeyt‘un]. *Yušarjan* 1911. 319–323.
- Malxasyanc‘, Step‘an. 1958. *Arakner avandut‘yunner anekdotner (Axalc‘xayi barbarov)* [Proverbs, traditional stories and anecdotes (in the dialect of Axalc‘xa)]. Yerevan: Academy Press.
- Margaryan, Alek‘sandr S. 1975. *Gorisi barbarə* [The dialect of Goris]. Yerevan: University Press.
- Markosyan, Րազմիկ Արարատի. 1989. *Araratyan barbar* [The dialect of Ararat]. Yerevan: “Luys”.
- Martirosyan, Hrach Kayosiki. 2010. *Etymological dictionary of the Armenian inherited lexicon*. Leiden & Boston: Brill. (Leiden Indo-European Etymological Dictionary Series 8).
- Martirosyan, Hrach Kayosiki. 2013. Armj anskie dialekty: xarakteristika otdel’nyx dialektov [Armenian dialects: Description of individual dialects]. In Yuri B. Koryakov & Andrej A. Kibrik (eds.), *Languages of the World: Relict Indo-European languages of Western and Central Asia*, 334–385. Moscow: Academia.
- Maxudianz, Mesrob. 1911. *Le parler arménien d’Akn* (quartier bas). Paris: Librairie Paul Geuthner.
- Meillet, A. 1936. *Esquisse d’une grammaire comparée de l’arménien classique*. Wien: Melchitharisten.
- Melik‘-Şahnazareanc‘, Kostandin K. (Tmblač‘i Xač‘an). 1907–1908 (1–2). *Zuřna-tmbla*. Vol. 1, 1907; vol. 2, 1908. Vâlarşapat: Elék‘trašarž Tparan M. At‘oroy S. Ējmiačni.
- Melik‘ean, Etišē Y. 1964. *Hark‘-Xnus*. Ant‘ilias, Lebanon: Tparan Kat‘ołikosut‘ean Hayoc‘ Meci Tann Kilikoy.
- Mežunc‘, B. X. 1989. *Šamšadin-Dilijani xosvack‘ə* [The subdialect of Šamšadin-Dilijan]. Yerevan: University Press.
- Mkrtč‘yan, Garik. 2015. *Barbaragtit‘yan himunk‘ner* [Foundations of dialectology]. (Lectures). Yerevan: “Daso Print”.
- Mkrtč‘yan, H. M. 1952. *Karno barbarə* [The dialect of Karin]. Yerevan: Academy Press.
- Mkrtč‘yan, Nerses Alek‘sani. 1971. *Burduri barbarə* [The dialect of Burdur]. Yerevan: Academy Press.
- Mkrtč‘yan, Nerses Alek‘sani. 1995. Sivri-Hisari gałt‘ðjaxi ew barbari masin [On the community and dialect of Sivri-Hisar]. *PBH* 1995, Nr 2. 205–214.
- Mkrtč‘yan, Nerses Alek‘sani. 2006. *Anatoliyai norahayt hay barbarnerə ew banahyusut‘yunə* [The newly-found Armenian dialects and folklore of Anatolia]. Yerevan: “Orenk‘ ew irakanut‘yun”. (Heleinakayin hratarakut‘yun).
- Movsisyan, Sahak (Bense). 1972. *Hark‘* (Mšo Bulanəx). *Hay azgagrut‘yun ew banahyusut‘yun* 3. Yerevan: Academy Press.
- Msereanc‘, Lewon. 1897–01. (Mserianc), *Ètjudy po armjanskoy dialektologii*, 1–2 (on the dialect of Muš). Moscow: University Press.
- Msereanc‘, Lewon. 1899. *Hayerēn barbaraxōsut‘iwn* [Armenian dialectology]. (Armenian transl. from Russian by Gabriēl Mēnēvišean). Vienna: Mxit‘arean Tparan. (Azgayin Matenadaran, 30).
- Muradean, Y. 1901. *Hamšēnc‘i hayer* [Hamšen Armenians]. T‘iflis: K. Martiroseanc‘i tparan (reprinted from *Azgagrakan handēs*).

- Muradyan, H. D. 1960. *Karčewani barbarə* [The dialect of Karčewan]. Yerevan: Academy Press.
- Muradyan, H. D. 1967. *Kak‘avaberdi barbarə* [The dialect of Kak‘avaberd]. Yerevan: Academy Press.
- Muradyan, H. D. 1982. *Hayoc‘ lezvi patmakan k‘erakanut‘yun*, vol. 1. *Hnč‘yunabanut‘yun* [Historical grammar of Armenian: Phonology]. Yerevan: Academy Press.
- Muradyan, M. H. 1962. *Šataxi barbarə* [The dialect of Šatax]. Yerevan: University Press.
- Muradyan, M. H. 1975. *Bařagitus‘yun* [Lexicology]. In *Hayoc‘ lezvi patmakan k‘erakanut‘yun*, vol. 2, 5–158. Yerevan: Academy Press.
- Muradyan, M. H. 1982. Urvagic Moksi barbaři [An outline of the dialect of Moks]. In *Hayereni barbaragitan atlas: usumnasirut‘yunner ew nyut‘er* 1, 108–183. Yerevan: Academy Press.
- Mxit‘areanc‘, A. 1901. *P‘šrank‘ Širaki ambarneric‘* [Fragment from the storehouses of Širak]. (Ēminean azgagrakan žołovacu, vol. 1). Moscow, Alek‘sandrapol.
- Nawasardeanc‘, Tigran. 1903. *Bařgirk‘ Araratean barbaři* [Dictionary of the dialect of Ararat]. T‘iflis: Tparan Movsēs Vardaneanc‘i.
- Nždehean, K‘. 1899. Alaškerti banawor grakanut‘iwn [Oral literature of Alaškert]. *Azgagrakan handēs* 4th year, 5th book, 1899, Nr 1. 185–199.
- Nždehean, K‘. 1902. Pařawašunč‘. *Azgagrakan handēs* 7th year, 9th book. 263–271.
- Nždehean, K‘. 1908. Alaškerti banawor grakanut‘iwnic‘ [From the oral literature of Alaškert]. *Azgagrakan handēs* 11th year, 18th book, 1908, Nr 2. 33–36.
- Nždehean, K‘. 1910. Alaškerti banawor grakanut‘iwnic‘ [From the oral literature of Alaškert]. *Azgagrakan handēs* 12th year, 20th book, 1910, Nr 2. 108–111.
- Orbeli, I. A. 2002. *Izbrannye trudy v dvux tomax* [Selected writings in two volumes], vol. 2. Yerevan: Academy Press “Gitutjun”.
- Őtean-Gasbarean, Alis. 1968. *Patmagirk‘ Ankiwrioy ew Stanozi hayoc‘* [Chronicle of the Armenians of Ankiwria and Stanoz]. Pěyrut‘: Tp. Tōnikean. [Alice Odian-Kasprian, The history of the Armenians of Angora and Stanos. Beirut: Doniguan Press].
- Patkanov, K. 1869. *Izsledovanie o dialektax armjanskago jazyka: filologičeskij opyt* [Study on Armenian dialects]. Sankt-Peterburg: Tipografija Imperatorskoj Akademii Nauk.
- Patkanov, K. 1875. *Materialy dlja izučenija armjanskix narečij* [Materials for the study of Armenian dialects]. Vol. 1: Govor Naxičevanskij; vol. 2: Mušskij dialekt. Sanktpeterburg: Tipografija Imperatorskoj Akademii Nauk.
- Petermann, H. 1867a. Über den armenischen Dialect von Agulis nach einer Mittheilung des Herrn v. Patkanoff (Patkanean). *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*: Sitzung der philosophisch-historischen Klasse vom 26 Nov. 1866. 727–741.
- Petermann, H. 1867b. Über den Dialect der Armenier von Tiflis. *Philologische und historische Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin* 1866 (gelesen am 16. Mai 1861). 57–87.
- Petoyan, Vardan A. 1954. *Sasuni barbarə* [The dialect of Sasun]. Yerevan: Academy Press.
- Petoyan, Vardan A. 1965. *Sasna azgagrut‘yunə* [The ethnography of Sasun]. Yerevan: Academy Press.
- Pisowicz, Andrzej. 1997. Consonant shifts in Armenian dialects during the post-Classical period revisited. In N. Awde (ed.), *Armenian perspectives*, 215–230. Richmond & Surrey: Curzon.
- Pisowicz, Andrzej. 2003. “Muslim” words in the Polish subdialect of the Ardeal dialect of Armenian. *Annual of Armenian linguistics* 22–23. 29–39.

- P'olataean, Andranik L. 1969. *Patmut 'iwn hayoc ' Arabkiri* [History of the Armenians of Arabkir]. New York: Baykar Press (Arabkir Union).
- Pōłosean, Y. P. 1942. *Hačəni əndhanur patmut 'iwnə ew šrjakay Gōzan-Tağı giwlerə* [General history of Hačən]. Los Angeles: Bozart Press (Pōzačean).
- Połosyan, A. M. 1965. *Hadrut i barbarə*. Yerevan: Academy Press.
- P'ork'şeyan, X. A. 1971. Nor Naxijewani hay žolovrdakan banahyusut'yunə [The folklore of Nor Naxijewan]. *Hay azgagrut'yun ew banahyusut'yun 2*. Yerevan: Academy Press.
- PtmSivHisHay. 1965. *Grigor Tēr Yovhannēsan* (ed.), *Patmagirk ' Sivri-Hisari hayoc ': tełagrakan, patmakan ew azgagrakan* [Chronicle of the Armenians of Sivri-Hisar: topographical, historical and ethnographical], (initiated by Mihran Nersēsan). Pēyrut': Tparan Mšak.
- Šahpazean, Y. 1913. *Bun haykakan stelcagorcut'yun* [Genuine Armenian creation]. Łalat'ia: Tparan "Šant".
- Sarafian, Gēorg A. 1953. *Patmut 'iwn Ant'ēpi hayoc '* [Armenian history of Ayntab], vol. 2. Los Angeles, California.
- Sargisean, Yarut'iwn K'hn. (Alewor). 1932. *Balu: ir sovoroyt'nerə, krt'akan u imac'akan vičakə ew barbarə* [Balu: its customs, educational and intellectual state and the dialect]. Cairo: "Sahak-Mesrop".
- Sargseanc', Sargsis. 1883. *Agulec'oc ' barbarə (zōkeri lezun): lezuabanakan hetazōtut 'iwn* [The dialect of Agulis: a linguistic study]. Vols. 1–2. Moscow: Skoropečatnja O. Gerbeka.
- Sargsyan, Armen Yuriki. 2013. *Łarabati barbarı bařaran* [Dictionary of the dialect of Łarabat]. Yerevan: "Edit Print".
- Sedrakean, Aristakēs. 1874. *K'nar mšec 'woc ' ew vanec 'woc '(minč'ew ayžm č 'hratarakuac erger)* [Lyre of inhabitants of Muš and Van: unpublished songs]. Vałarşapat: Tparan Srboy Kat'ulikē Ĕjmiacni.
- Şērenc', Gēorg G. 1885–99. (VanSaz 1–2), *Vanay saz*. Vol. 1. Hawak'acoyk' Vaspurakani žolovrdakan ergeri, hēk'eat'neri, ařacneri ew hanelukneri (tełakan barbařov). T'iiflis: Tparan Yovhannēs Martiroseanc'i, 1885. Vol. 2. Hawak'acoyk' Vaspurakani žolovrdakan ergeri, hēk'eat'neri, ařacneri, hanelukneri, örhnut'iwnneri ew anēck'neri (tełakan barbařov). T'iiflis: Tparan K. Martiroseanc'i, 1899.
- Simonian, Hovann H. (ed.). 2007. *The Hemshin: History, society and identity in the Highlands of Northeast Turkey*. London & New York: Routledge.
- Simonyan, N. M. 1979. Hay barbařneri hndevropakan hnabanut'yunnerə [Indo-European archaisms of the Armenian dialects]. In *Hayoc ' lezvi hamematakan k'erakanut'yan harc 'er*, 188–248. Yerevan: Academy Press.
- Srapean, Lewon. 1960. K'li: *Tełagrakan ew azgagrakan* [K'li: topography and ethnography]. Ant'ilias, Lebanon: Tparan Kat'olikosut'ean Hayoc' Meci Tann Kilikioy.
- Sruanjteanc', Garegin. 1978–1982. (Sravanjtyanc') *Erker*. Vol. 1: 1978; vol. 2: 1982. Yerevan: Academy Press.
- Svazlyan, Veržine Garñiki. 1984. Musa ler. *Hay azgagrut'yun ew banahyusut'yun 16*. Yerevan: Academy Press.
- Svazlyan, Veržine Garñiki. 1994. *Kilikia: Arewmtahayoc ' banavor avandut'yunə* [Cilicia: The oral tradition of Western Armenians]. Yerevan: Academy Press "Gitut'yun".
- Svazlyan, Veržine Garñiki. 2000. *Polsahayoc ' banahiwsut'iwnə* [Folklore of the Armenians of Polis]. Yerevan: Academy Press "Gitut'yun".

- Tarōnean, Eduard. 1961. *Bałeši nahangi Xult‘ik giwli barbarə* [The dialect of the Xult‘ik village in the province of Bałeš]. Ant‘ilias, Lebanon: Tparan Kat‘olikosut‘ean Hayoc‘ Meci Tann Kilikoy.
- T‘emurčyan, V. S. 1970. *Gamirk‘i hayerə* [The Armenians of Gamirk‘]. *Hay azgagrut‘yun ew banahusut‘yun* 1. 7–140.
- Tēovlēt‘ean, Sargis X. 1954. *Yušer Šapin Garahisarēn* [Memories from Šapin Garahisar]. P‘ariz: “Le soleil”.
- Tēr-Alek‘sandrean, Gēorg. 1885. *T‘iflisec‘oc‘ mtawor keank‘ə* [Intellectual life of inhabitants of T‘iflis]. T‘iflis: Yovhannēs Martiroseanc‘i tparan.
- Tēr-Pōlosean, Gr. 1921–22. *Havariki barbarə* [The dialect of Havarik]. *Banber Hayastani Gitakan Instituti*. Vol. 1–2. 160–177.
- Tēr-Sargsenc‘, Aristakēs vardapet. 1875. *Panduxt vanc‘in* [The emigrant from Van]. K. Pōlis: Tntesean.
- Tēr-Vardanean, Gurgēn. 1968. *Erznka-Kamax gawařabarbařə ew azgagrakan yušer* [The dialect of Erznka-Kamax and ethnographical memories. Erusalēm.
- Tēr-Yakobean, Yakob. 1960. *Partizakə xatutik. Yaweluacabar – Manušak ew ir hēk ‘eat’nerə*. Paris.
- Tomaschek, Wilhelm. 1896. Sasun und das Quellengebiet des Tigris: historisch-topographische Untersuchung. *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften*. 133.4. Wien.
- Tomson, A. 1887. *Lingvističeskaja izsledovanija* [Linguistic studies], vol. 1. S. Peterburg: Tipografija E. Evdokimova.
- Tomson, A. 1890. *Istoričeskaja grammatika sovremenennago armjanskago jazyka goroda Tiflisa* [Historical grammar of the modern Armenian language of the city of Tiflis]. S.-Peterburg: Tipografija E. Evdokimova.
- T‘ořlak‘yan, Barunak G. 1981. *Hamšenahayeri azgagrut‘yuna* [Ethnography of the Hamšen Armenians]. *Hay azgagrut‘yun ew banahusut‘yun* 13. Yerevan: Academy Press.
- T‘ořlak‘yan, Barunak G. 1986. *Nšxarner Hamšeni ew Trapizoni banahusut‘yan* [Fragments of folklore of Hamšen and Trapizon]. Yerevan: University Press.
- T‘umayean, Karapet. 1930. *Marzuanı gawařabarə* [The dialect of Marzuan]. Vienna: Mxit‘arean tparan.
- T‘uršyan, Harut‘yun. 2018. *Bararan Vani barbari* [Dictionary of Van dialect] (Dr. Haykanush Mesropyan, ed.). Yerevan: “Gitutyun” publishing house of NAS RA.
- Vardanyan, Sergey. 2009. *Kronap‘ox hamšenahayeri barbarə, banahusut‘yuna ew ergarvestə: nyut‘er ew usumnasirut‘yunner* [The dialect, folklore and song art of converted Armenians of Hamšen: materials and studies]. Yerevan: University Press.
- Vartapetjan, N. A. 1962. *Sjunikskij dialekt* [The dialect of Syunik‘]. (PhD dissertation). Yerevan: Academy Press.
- Vaux, Bert. 1992. Adjarian’s Law and consonantal ATR in Armenian. In *Proceedings of the Fourth International Conference on Armenian Linguistics, Cleveland State University, Cleveland, Ohio, September 14–18, 1991*, 271–293. (John A. C. Greppin, ed.). Delmar & New York: Caravan Books.
- Vaux, Bert. 1997. The phonology of voiced aspirates in the Armenian dialect of New Julfa. In N. Awde (ed.), *Armenian Perspectives*, 231–248. Richmond, Surrey: Curzon.
- Vaux, Bert. 1998. *The phonology of Armenian*. Oxford: Clarendon Press.
- Vaux, Bert. 1999–2000. Notes on the Armenian dialect of Ayntab. *Annual of Armenian linguistics* 20. 55–82.

- Vaux, Bert. 2000–2001. Hemshinli: The forgotten Black Sea Armenians. *Journal of Armenian studies* 6(2). 47–71.
- Vaux, Bert. 2001. The Armenian dialect of Aslanbeg. *Annual of Armenian linguistics* 21. 31–64.
- Vaux, Bert. 2007. Homshetsma: The language of the Armenians of Hamshen. In Hovann H. Simonian (ed.), *The Hemshin: History, society and identity in the Highlands of Northeast Turkey*, 257–278. London & New York: Routledge.
- Viredaz, Rémy. 2003. The Great Armenian Puzzle is not thorny. *Annual of Armenian linguistics* 22–23. 59–81.
- Weitenberg, Jos J. S. 1985. Additional -n in Armenian. *Annual of Armenian linguistics* 6. 101–106.
- Weitenberg, Jos J. S. 1986. Additional h-, initial y- and Indo-European *y- in Armenian. In Maurice Leroy & Francine Mawet (eds.), *La place de l'arménien dans les langues indo-européennes*, 90–101. Louvain: Peeters.
- Weitenberg, Jos J. S. 1992. An early dialectal phonetic split in Classical Armenian: the cases of *t'awt'ap'em* and *p'lanim*. In John A. C. Greppin (ed.), *Proceedings of the Fourth International Conference on Armenian Linguistics, Cleveland State University, Cleveland, Ohio, September 14–18, 1991*, 295–318. Delmar, New York: Caravan Books.
- Weitenberg, Jos J. S. 1993. Initial y- in Classical Armenian and the early development of Armenian dialects. In L. Hovsepian, N. Parnassian & S. Simonian (eds.), *Proceedings of the Second International Symposium on Armenian Linguistics, 21–23 September 1987*, 208–217. Yerevan: Armenian Academy Press.
- Weitenberg, Jos J. S. 1996. On the early development of the Armenian dialects. In Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University of Montreal, Quebec, Canada May 1–5, 1995*, 93–118. Delmar & New York: Caravan Books.
- Weitenberg, Jos J. S. 1999–2000. On the early development of Armenian dialects – 2. The monophthongization of *ay*. *Annual of Armenian linguistics* 20. 1–26.
- Weitenberg, Jos J. S. 2001. Thoughts on early Armenian accentuation. *Annual of Armenian linguistics* 21. 65–73.
- Weitenberg, Jos J. S. 2002. Aspects of Armenian dialectology. In Jan Berns & Jaap van Marle (eds.), *Present-day dialectology: Problems and findings*, 141–157. Berlin & New York: De Gruyter Mouton. (Trends in linguistics: Studies and monographs, 137).
- Winter, Werner. 1966. Traces of early dialectal diversity in Old Armenian. In H. Birnbaum & J. Puhvel (eds.), *Ancient Indo-European dialects: proceedings of the Conference on Indo-European linguistics held at the University of California, Los Angeles, 1963*, 201–211. Berkeley & Los Angeles: University of California Press.
- Xač'atryan, A. E. 1984. Jayndarji hin erewuyt'ner hay barbarnerum [Old phenomena of ablaut in Armenian dialects]. In *International symposium on Armenian linguistics (Yerevan, September 21–25, 1982): Reports*, 318–326. Yerevan: Academy Press.
- Xač'atryan, H. S. 1985. Šapin Garahisari barbar (holovuma ew xonarhum) [The dialect of Šapin Garahisar: declension and conjugation]. In *Hayereni barbaragitakan atlas: usumnasirut'yunner ew nyut'er 2*, 136–159. Yerevan: Academy Press.
- Xač'atryan, Raisa. 1999. T'alin. Hay azgagrut'yun ew banahyusut'yun 20. Yerevan: Academy Press “Gitut‘yun”.
- Xač'atryan, V. 2004. *Vardenisi (Diadini) barbarə* [The dialect of Vardenis (Diadin)]. Yerevan: “Nairi”.

- Xemč‘yan, Ėst‘er H. 2000. Tavuš. *Hay azgagrut‘yun ew banahyusut‘yun* 21. Yerevan: Academy Press “Gitut‘yun”.
- Xemč‘yan, Ėst‘er H. 2008. Ijewan (Jorop‘or). *Hay azgagrut‘yun ew banahyusut‘yun* 25. Yerevan: Academy Press “Gitut‘yun”.
- Yovakimean, Yovakim (Aršakuni). 1967. *Patmut‘iwn Haykakan Pontosi. Pēyrut‘: “Mšak”*. Yovnanean, Lewond. 1897. *Hetzazōtut‘iwnk‘ naxneac‘ ramkōrēni vray: usumnasirut‘iwnk‘ ew k‘atuack‘ner* [Studies on the dialect of ancients: studies and excerpts]. 2. vols. Venice: S. Lazar. (Azgayin Matenadaran, 23–24).
- Yovsēp‘eanc‘, Garegin Sarkawag. 1892. *P‘šrank‘ner žołovrdakan banahyusut‘iwnic‘* [Fragments from folklore]. T‘iflis: M. Šarajē.
- YušamXotorj. 1964. *Yušamatean Xotorjuri* [Monument of Xotorjur] (H. Yarut‘iwn v. Hulunean & H. Matt‘eos v. Hačeian, eds.). Wien: Mechitharisten.
- Zak‘aryan, Mari G. 2008. *Agulisi barbarə: ink‘nut‘yan bac‘ahaytum* [The dialect of Agulis: discovery of the identity]. Yerevan: “Iravunk”.

2.3. Northern Kurdish (Kurmanjî)

Geoffrey Haig

1. Introduction

This chapter presents an overview of Northern Kurdish, or Kurmanjî (in various spellings), as it is spoken in eastern Anatolia. The material is largely based on two joint publications with Ergin Öpentin: Öpentin and Haig (2014), and Haig and Öpentin (2018), to which the reader is referred for further details. The variety of Kurmanjî spoken in northern Iraq (Behdinî, under various spellings) is treated in Haig (this volume, chapter 3.3, §4). For the purposes of this chapter, “eastern Anatolia” is taken to coincide with the the eastern part of Turkey, extending south-eastward from a line beginning from Sivas, but excluding the Mediterranean and Black Sea coastal regions.

Eastern Anatolia is the homeland of the majority of Kurmanjî speakers, but their traditional settlement region overlaps into Syria, northern Iraq and western Iran (see Figure 1). Various estimates put the number of Kurmanjî speakers in Turkey at between 8 and 15 million, but any figures must be treated with caution, due to differences in definitions and methodologies. Following decades of violence in the region, there has been a large-scale diaspora from Anatolia’s rural areas to Turkey’s larger cities, both within Anatolia (e. g. Van, Diyarbakır, and Mardin) and in western Turkey (e. g. İstanbul, İzmir), and beyond into Europe. Perhaps as many as one third of Kurmanjî speakers have left their Anatolian homeland in the last 40 years, but reliable figures are impossible to obtain.

Traditionally, Kurmanjî is classified as a member of the northwest Iranian branch of the west Iranian languages, within the Iranian branch of Indo-European. Besides Kurmanjî, two other closely-related Iranian languages are spoken in Turkey, but are not covered in this volume. The first are so-called Şêx Bîzinî dialects, the language of the descendants of southern Kurdish tribes re-settled in various parts of Anatolia in the 16th century. With the exception of short descriptions in Lewendî (1997), which demonstrate beyond doubt the southern Kurdish origins of the dialects, further data on these varieties are unavailable to me. The second is Zazaki, spoken in several locations in central Anatolia (cf. Figure 1, and Paul 2009 for recent summary of Zazaki). Figure 1 illustrates the locations of the various varieties belonging to what is traditionally termed “Kurdish”. Just how one would define the assumed superordinate entity “Kurdish”, and which varieties should be included in it, are questions that go beyond the scope of this chapter; see Öpentin and Haig (2014), Haig and Öpentin (2018) for discussion.

Research on Kurdish in Turkey was hampered for decades by exclusory language politics directed at “non-Turkish” languages (see Haig 2004; Öpentin 2012, 2015;

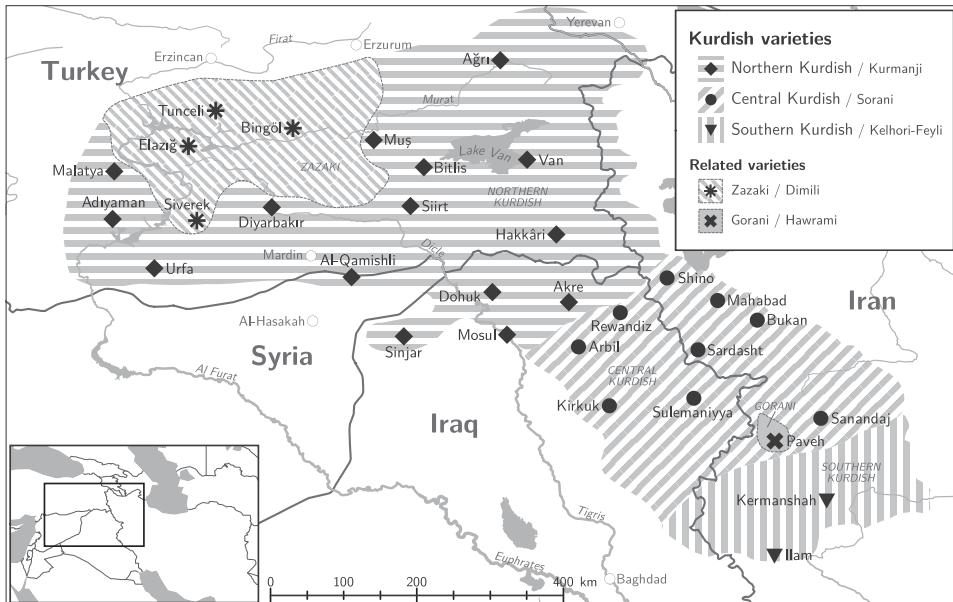


Figure 1: Overview of varieties traditionally considered to be “Kurdish”

Haig and Öpengin (2018) for discussion and references). Between around 1920 and 1990, the only reliable publications on Kurdish spoken in Turkey are Jastrow’s sketch of the phonology of the Van dialect (1977), and Ritter’s collection of spoken Kurdish narratives from the Midyat region (1971 and 1976). A rich body of material has also been compiled under the auspices of L’Institut kurde de Paris, though we still await a more structured survey. Very recently, an online-accessible data-base of Kurdish dialects has been launched (Matras et al. 2016), which provides the most comprehensive coverage to date of the Kurmanjî varieties spoken in eastern Anatolia. Although there is considerable dialectal variation within Anatolia, the situation is one of a dialect continuum, with a high degree of mutual intelligibility across most of the region, except between the farthest southeastern and northwestern dialects (see §4 below). In this regard, the situation of Kurmanjî is rather different to that of Neo-Aramaic, likewise spoken in eastern Anatolia and northern Iraq, but which consists of scattered and distinct local dialects (or perhaps languages), with low levels of mutual intelligibility (see Khan, this volume, chapters 2.5 and 3.4).

Below I provide a brief outline of the history and current situation of the speakers (Section 1.2), followed by a grammatical sketch of what is termed here “Standard Kurmanjî”, loosely based on the dialect of Cizre and Botan in southeastern Turkey (Section 2). In Section three, selected areas of variation are presented, illustrating the two most divergent dialects in eastern Anatolia, namely southeast-

ern Kurmanjî and western Kurmanjî. Section four sums up the main points of the chapter.

1.1. History of the Kurmanjî speech community in Anatolia

There are no direct historical records of Kurdish settlement in Anatolia, so estimates of the timing and pathways are correspondingly speculative. Asatrian (2009) follows MacKenzie (1961b) in assuming that the initial formation of Kurdish would have occurred in a southwest Iranian environment, namely the northern areas of Fars in Iran. Northern expansion of the Kurds into what was then Armenia began, according to Asatrian, based on the first attestations of Kurds in Armenian texts, between the 8–9 centuries AD and continued over a period of several centuries. Some clues regarding the chronology of these events can be obtained from the dating of Armenian loan words in Kurdish. The point of origin of this northward expansion is considered to be “Northern Iraq, Hakkari, southern shore of Lake Van” (Asatrian 2009: 35).

Haig and Öpengin (2018) suggest that a northwestward expansion of the Kurmanjî-speaking peoples into a largely Armenian-populated region receives some support from the areal distribution of certain features in Kurmanjî. The first is the comparative homogeneity, and relative simplicity (in terms of morphology) of those Kurmanjî dialects further to the west. The southeastern varieties of Kurmanjî (cf. e. g. Behdinî in North Iraq (Haig, this volume, chapter 3.3, §4) and Şemzinan in Turkey, show the richest morphology, the most complex syllable structures, the most consistent retention of gender and ergativity, all items that can be considered typically “mature features” (Dahl 2004; Trudgill 2011), of the kind which reflect a comparatively long period of stable settlement. The other varieties, on the other hand, all show, to varying degrees, loss of these features, which would be expected under conditions of mobility and language contact involved in the northwestward expansion of the Kurds (and possible shift to Kurdish among speakers of other languages such as Armenian, Neo-Aramaic or Arabic, cf. Trudgill 2011, and McWhorter 2005 for the role of contact in simplification). At any rate, there must have been a strong Armenian influence on Kurmanjî in these formative stages, evidenced in the presence of an additional row of unaspirated voiceless plosives and affricates in Kurmanjî, a feature that is characteristic (to varying degrees) of all the dialects (cf. Section 3.1).

Several thousand Kurmanjî speakers were forcibly resettled in the Khorasan province of Iran in the seventeenth century by the Safavids, with the intention of protecting the eastern borders of Persia from Uzbek and Turkmen tribes (Oberling 2008). They have maintained their language and many customs down to the present. Although there is little systematic research on their language, my impression of recordings made available to me by Don Stilo (p.c.) is that they are descendants of speakers from the western Kurmanjî dialect group (see §3.5 below).

In Anatolia, Kurmanjî speakers have co-existed with speakers of Armenian, Neo-Aramaic, Zazaki, Arabic and Turkish for many centuries, leading to considerable mutual influence. In the southeast, Kurdish has left a deep impact on Arabic and Neo-Aramaic (see Khan, this volume, chapters 2.5 and 3.4), while Armenian left its mark at least in the phonology (see §3.1). Turkish influence on Kurmanjî has also been considerable (Haig 2001, 2006, 2007), particularly on Western Kurmanjî (see §4). However, it is probably first since the founding of the Turkish Republic in 1923, leading to nation-wide compulsory schooling and military service, and the large-scale diaspora from rural Anatolia in the second half of the twentieth century, that Turkish influence on Kurmanjî has become so omnipresent (see §4). The future of Kurmanjî in Turkey is impossible to predict. But it must be noted that, despite the seemingly high absolute figures of “Kurmanjî speakers” today, very few children are exposed to the early rich input that would ensure full acquisition of the language, and there is still no regular representation of Kurmanjî in state-controlled primary education.

2. A grammatical sketch of “Standard Kurmanjî”

Although Kurmanjî is spoken across a large area, there is a reasonably widely-accepted written standard, loosely based on the dialect of the region Cizre and Botan in southeastern Turkey. It uses a modified version of the Roman alphabet, and is employed in all manner of publications, including journals, newspapers, literature, internet publications, chat-rooms etc. Today’s norms are largely based on the standards established by Celadet Ali Bedir Khan in a series of articles in the journal *Hawar*, published in the 1930’s. These conventions were later codified in Bedir-Khan and Lescot’s *Grammaire kurde (dialecte kurmandji)*, which was published in 1970. In this chapter I will present a grammatical sketch of this Standard Kurmanjî (Standard K.), based on Haig and Öpentin (2018, §3), with examples provided in standard orthography, but noting additional phonetic detail where necessary. In Section 3, selected issues in regional variation are presented.

2.1. Phonology

2.1.1. Vowels

The basic vowel system consists of eight simple vowel phonemes, five long (or tense) vowels, and three short, or lax vowels.

Long, or tense vowels: /a/ <a>; /i/ <î>; /e/ <ê>; /o/ <o>; /u/ <û>

Short, or lax vowels: /æ/ <e>, /ʊ/ <u>, /ɪ/ <i>

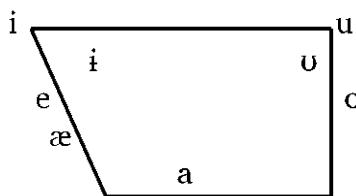


Figure 2: Pan-dialectal scheme for the vowel phonemes of Kurdish
(cf. Haig and Öpenguin, 2018)

The tense vowels are /i/, /e/, /a/, /u/ and /o/. They are generally realized phonetically long, particularly in open syllables, and indeed, they are the ones that are stretched in traditional Kurdish songs. However, vowel length by itself is not phonemically distinctive in Kurdish. The full vowels occupy approximately the five positions of a fairly typical five-term vowel system; the other three vowels are more centralized. Examples from Haig and Öpenguin (2018, §3.1.1) are presented in Table 1.

Table 1: The tense vowels of Kurmanjî (from Haig and Öpenguin, 2018)

/a:/ <a>			/e:/ <ê>		
Orthography	IPA	Gloss	Orthography	IPA	Gloss
<i>agir</i>	[a:gir]	fire	<i>êvar</i>	[e:var]	evening
<i>sar</i>	[sa:r]	cold	<i>sêv</i>	[se:v]	apple
<i>mal</i>	[ma:l]	house, home	<i>lêv</i>	[le:v]	lip
<i>mar</i>	[ma:r]	snake	<i>t'êr</i>	[te:r]	satiated
<i>zava</i>	[za:va:]	bridegroom	<i>avê</i>	[ave:]	water(OBL)

/i:/ <î>			/u:/ <û>		
Orthography	IPA	Gloss	Orthography	IPA	Gloss
<i>spî</i>	[s ^h pi:]	white	<i>bûk</i>	[bu:k]	bride
<i>bîne</i>	[bi:næ]	bring!	<i>zû</i>	[zu:]	soon, fast
<i>tari</i>	[tari:]	dark(ness)	<i>çû</i>	[ç ^h u:]	s/he went
<i>nîne</i>	[ni:næ]	there isn't	<i>xwesû</i>	[x ^w æsu:]	mother-in-law
<i>dîk</i>	[di:k]	rooster	<i>tûj</i>	[tu: ^h ʒ]	sharp

/o/ <o>		
Orthography	IPA	Gloss
<i>toz</i>	[t ^h oz]	dust
<i>çok</i>	[t ^h ok]	knee
<i>got</i>	[got]	said
<i>zozan</i>	[zo:zan]	alpine summer settlement
<i>koçer</i>	[koç ^h ær]	nomad

The two lax vowels /œ/ and /æ/ are realized more centrally than the tense vowels. They are less prone to lengthening in open syllables, but are not subject to elision under the phonological processes to be discussed below. They may also occur at the end of words. Examples are provided in Table 2:

Table 2: The lax vowels of Kurmanji (from Haig and Öpentin, 2018)

/æ/ <e>			/œ/ <u>		
Orthography	IPA	Gloss	Orthography	IPA	Gloss
<i>em</i>	[æm]	we	<i>guh</i>	[gu(h)]	ear
<i>dest</i>	[dæst]	hand	<i>kuştin</i>	[kuştin]	kill
<i>ser</i>	[sær]	head	<i>gund</i>	[gond]	village
<i>dev</i>	[dæv]	mouth	<i>quling</i>	[qolinq]	crane (bird)
<i>re</i>	[ræf]	black	<i>xurt</i>	[xort]	strong, sturdy

The mid-high, centralized rounded vowel /œ/ is sometimes difficult to distinguish from the mid-high, unrounded /i/, leading to variation across dialects and in the spelling of some words, e. g. *muhacir* ~ *mihacir* ‘refugee’, *tucar* ~ *ticar* ‘trader’.

The central vowel /i/ is approximately a mid-high, mid-closed, unrounded vowel; it cannot occur word-finally (though there are dialectal exceptions; see §3.1). Haig and Öpentin (2018) distinguish two underlying sources of this vowel, though both are written with the same symbol <i> in the standard orthography: the lexical central vowel, and the epenthetic central vowel. The lexical version occurs as a stressed vowel in the stems of lexical items; it is not subject to the deletion processes that affect the epenthetic central vowel.¹ Some examples are provided in Table 3.

Table 3: The lexical central vowel in selected words

Orthography	IPA	Gloss
<i>mij</i>	[miz]	fog, mist
<i>pirç</i>	[pʰirç]	hair (of head)
<i>dil</i>	[dil]	heart
<i>diz</i>	[diz]	thief
<i>kir</i>	[kif]	did

Epenthetic central vowels occur in a number of forms, generally in order to satisfy constraints on syllable structure. An example is the use of epenthetic vowels in the infinitive endings of verbs. If we assume that the infinitive ending is [-n], then this

¹ When the surrounding consonants are sonorants, as in *dimirim* ‘I die’, a lexical vowel may be elided.

may attach directly to a vowel-final past verb stem, such as *kēṣa-* ‘draw, pull’ (note that some dialects have a different past stem for this verb). The infinitive is thus *kēṣa-n*. Following a consonant-final stem, however, an epenthetic [i] is inserted to avoid non-licensed syllable codas; see Table 4 for examples

Table 4: Epenthetic vowels in infinitives

Past stem	Infinitive	Gloss
<i>kēṣa-</i>	<i>kēṣa-n</i>	pull, smoke (cigarettes)
<i>hat-</i>	<i>hat-in</i>	come
<i>dīt-</i>	<i>dīt-in</i>	see
<i>nasī-</i>	<i>nasī-n</i>	know (a person)
<i>xwend-</i>	<i>xwend-in</i>	read, study, recite

Vowel epenthesis may also occur in syllable-onset clusters in lexical items, though there is considerable regional variation here:

<i>bilind</i> ~ <i>blind</i>	‘high’	<i>sitra:n</i> ~ <i>stra:n</i>	‘song’
<i>bilu:r</i> ~ <i>blu:r</i>	‘type of wooden flute’	<i>bira: ~ bra:</i>	‘brother’
<i>dire:ʒ</i> ~ <i>dre:ʒ</i>	‘long’	<i>fɪkæft ~ fɪkæft</i>	‘cave’
<i>firotin</i> ~ <i>frotin</i>	‘sell’	<i>zima:n</i> ~ <i>zma:n</i>	‘tongue, language’
<i>sipi: ~ spi:</i>	‘white’	<i>zila:m</i> ~ <i>zla:m</i>	‘man’

Certain inflectional prefixes, and prepositions, consist of a single consonant plus the central vowel /i/. Examples are:

<i>di-</i>	Indicative present
<i>bi-</i>	Subjunctive
<i>li</i>	‘at, in’
<i>bi</i>	‘through’
<i>ji</i>	‘from’

In these formatives, the central vowel can also be considered epenthetic rather than lexical. It tends to be deleted under certain conditions (and in some dialects these vowels are seldom realized, regardless of phonological conditioning). For example the preposition *ji* [ʒi] ‘from’ undergoes reduction and devoicing in normal connected speech: ‘from you’ [ʒi tæ > stæ].

2.1.2. Consonants

The consonant phonemes of Kurmanjî are shown in Table 5.

Table 5: The consonant phonemes of Kurmanjî, generalized scheme (Haig and Öpengin, 2018)

	bilab.	lab.- dent.	dent	alveol	post-alv.	pal.	vel.	uvul.	pharyn.	glott.
Plos.	p ^h p b			t ^h t d		k ^h k g	q		?	
Fric.		v f		ʃ ʒ		x ɣ	ʁ	ħ ʕ ²	h	
Affr.				ʃ̡ ʒ̡						
Nas.	m			n			ŋ			
Trill				r						
Flap				r̡						
Approx.	w					j				
Lateral			l (dialectally also t̡)							

The most unusual feature of the consonant system is the three-way contrast on the stops and affricates, which emerged most probably through Armenian influence, illustrated with examples in Table 6.

Table 6: Three-way contrast on the stops and affricates (Haig and Öpengin, 2018)

voiceless aspirated:	[p ^h o:r]	'hair'	[t ^h æv]	'together'
voiceless, unaspirated:	[po:z]	'nose'	[tævir]	'hoe, mattock'
voiced:	[bo:z]	'grey-white (of horses)'	[dæv]	'mouth'
voiceless aspirated:	[k ^h a:r]	'work, matter, concern'	[tʃ ^h ima:]	'why'
voiceless, unaspirated:	[ka:t̡]	'old man'	[tʃæm]	'stream, brook'
voiced:	[ga:v]	'step, time'	[dʒæm]	'by, beside'

² The phonemic status of the pharyngeal sounds in Kurmanjî is controversial. First, they are most prominently linked to Semitic loan words, though Barry (2017) points to language-internal factors that have contributed to the emergence of pharyngeals, and their spread to native vocabulary (e. g. most dialects have initial [ħ] in the word for the numeral “7”). Second, the extent to which they are realized is subject to considerable cross-dialectal variation. Finally, as pointed out by Christiane Bulut (p.c.), in Kurdish as well as other languages of the region, the corresponding segments can be considered to be glottal stops produced with a retracted tongue root, rather than fricatives. Given their prominence in at least some varieties, we include them in Table 2. We also note that pharyngealization may be a feature that permeates over an entire syllable, rather than being localizable on a single segment.

All word-initial <r> sounds are trilled, but in other environments the distribution is not predictable. Examples for trilled and flap <r> are as follows:

Trilled	Flap
[pir] ‘much, many’	[pir] ‘bridge’
[kær] ‘deaf’	[kʰær] ‘donkey’
[biri:n] ‘to cut’	[biri:n] ‘wound’

With regard to pharyngeal segments, there is considerable cross-dialect variation, (see Khan 2008 on pharyngealization as a variant feature of pronunciation, and Haig and Öpengin, 2018, §4.2.1 for discussion of local variation). Some relatively widespread examples include [ɿeli:] ‘Ali’; [teɻm] ‘taste’; [pʰeħn] ‘flat’.

2.2. Nominal morphology

2.2.1. Gender and case

Nouns have an inherent two-way gender distinction between masculine and feminine. The difference is reflected formally in the form of the *ezafe*, and in the form of the singular *Oblique* case marker. In the plural, all gender distinctions are neutralized. Gender assignment is partially semantically motivated: words that refer to human beings and higher animals with a particular sex, such as *mehîn* ‘mare’ and *ap* ‘uncle’ are assigned grammatical gender according to their biological sex. Words that refer to persons, but which are usable with reference to either sex (e. g. *heval* ‘friend’) have no lexically fixed gender. Gender assignment with such words is determined according to the intended reference in a given context (Haig and Öpengin 2015).

For words denoting inanimate objects, or smaller animals, the principles of gender assignment are opaque. There are some morphological regularities accounting for gender, for example nouns created with the derivational suffix *-(y)i* are feminine, as are the infinitives of all verbs. In the dialects of Turkey, with the exception of those close to Behdinī, the default gender for inanimate nouns is feminine: most loanwords with non-human reference take this gender. In Behdinī, on the other hand, the default gender is masculine. Dialectal variation in gender is discussed in Haig and Öpengin (2018, §3.2.2).

There is a two-way case distinction between Direct (unmarked) and Oblique. Indefiniteness is marked on singular nouns through the suffix *-ek*, while no dedicated definiteness marker exists. A bare noun may thus have either a singular, definite reading, or a generic, sortal reading, depending on the context. Paradigms for singular nouns showing their inflectional possibilities are shown in Table 7.

Table 7: Case and indefiniteness in Standard Kurmanjî (Haig and Öpentin, 2018)

Masculine singular				Feminine singular				
Definite		Indefinite		Definite		Indefinite		
Dir.	Obl.	Dir.	Obl.	Dir.	Obl.	Dir.	Obl.	
<i>gund</i>	<i>gund-i</i>	<i>gund-ek</i>	<i>gund-ek-i</i>	<i>jin</i>		<i>jin-ê</i>	<i>jin-ek</i>	<i>jin-ek-ê</i>

The forms for oblique case marking on singular nouns are provided in Table 7. Note that the expression of the oblique case is suppressed when the noun concerned is followed by the *ezafe*, and it may be absent with singular masculine nouns (see below). The oblique case is used in the following syntactic functions:

- I. Object of a present-tense transitive verb
- II. Goal or Recipient argument immediately following a predicate of motion or transfer
- III. Complement of any adposition (though dropped in certain combinations)
- IV. Possessor in an *ezafe* construction
- V. Subject of a past-tense transitive verb

The direct case is used elsewhere.

Zero is common for proper nouns, and for masculine singular nouns particularly when they have generic reference, in most dialects of Central Anatolia. The following example is from the Kurdish textbook *Hînker*:

- (1) Ez şîr ve-na-xw-im
 I milk(M) PRV-IND-drink.PRS-1SG
 ‘I do not drink milk.’

Stem-vowel raising is found in many dialects; it only affects the open, non-rounded vowels [a, æ], when they are in stressed syllables, and raises them: [a, æ → e:]. For example:

<i>aş</i>	<i>êş</i> ‘mill’	<i>ga</i>	<i>gê</i> ‘ox’
<i>nan</i>	<i>nên</i> ‘bread’	<i>ba</i>	<i>bê</i> ‘wind’
<i>baxçe</i>	<i>bêxçe</i> ‘garden’	<i>kevir</i>	<i>kêvir</i> ‘stone’
<i>bajar</i>	<i>bajér</i> ³ ‘town’	<i>zilam</i>	<i>zilêm</i> ‘man’
<i>hesp</i>	<i>hêsp</i> ‘horse’	<i>xanî</i>	<i>xêni</i> ‘house’
<i>şivan</i>	<i>şivên</i> ‘shepherd, goatherd’	<i>lawik</i>	<i>lêwik</i> ‘boy, son’
<i>welat</i>	<i>welêt</i> ‘state, country’	<i>ezman</i>	<i>ezmên</i> ‘sky’
<i>ziman</i>	<i>zimên</i> ‘tongue, language’		

³ In some dialects where stem-vowel raising is not an option for marking the oblique case, the raised form *bajér* ‘town’ has become the unmarked form of the noun, used in all contexts, implying that the rule existed at earlier stages of the language.

Bare masculine singular nouns only consistently receive a suffixal marking of the oblique in the Badinan dialects of North Iraq, and in the east of the Hekari region in Turkey. As noted above, suffixation is regularly and consistently applied to all masculine singular nouns, and across all dialects, when the NP concerned has a determiner such as a demonstrative, or the interrogative *kîjan* ‘which?’, or carries the indefiniteness suffix *-ek*. An example with a demonstrative is (2); the presence of an oblique suffix suppresses stem-vowel raising:

- (2) *li vî welat-î ...* (not: **li vî welêt-î ...* or **li vî welêt ...*)
 in DEM.OBL homeland-OBL.M
 ‘in this homeland’

See Haig and Öpengin (2018, §3.2.4) for regional variation in the marking of singular masculine obliques.

2.2.2. Plural number

In Standard K., only nouns in the oblique case are overtly marked for plural, through the suffix *-a(n)* (deletion of *-n* is normal in some dialects), as shown in Table 8:

Table 8: Plural and case marking in Standard Kurdish

Plural (masc. and fem.)			
Definite		Indefinite	
DIRECT	OBLIQUE	DIRECT	OBLIQUE
<i>jin</i>	<i>jin-a(n)</i>	<i>jin(-in)</i>	<i>jin-a(n)</i>
<i>gund</i>	<i>gund-a(n)</i>	<i>gund(-in)</i>	<i>gund-a(n)</i>

An ending for indefinite direct plural *-in* is regularly cited in pedagogical works and is shown in brackets above, but it is only frequently attested in the dialects of Mardin region, and across the border in Syria. Elsewhere it is rare or lacking completely.

Nouns in the direct case do not inflect for plural. Such nouns are usually subjects, so plurality is generally reflected in number agreement on the verb:

- zarok hat-in* ‘the children came’
zarok hat ‘the child came’

There is a tendency in the dialects to the west for the Oblique plural marker to become a general plural marker, which is used on nouns in the direct case, and also on the demonstratives; see below on Mereš dialect.

2.2.3. Pronouns

The forms for the personal pronouns are given in Table 9. The “third person” pronouns are basically the distal demonstratives.

Table 9: Personal pronouns in Standard Kurmanjî (cf. Haig and Öpentin 2018)

		DIRECT	OBLIQUE
SG	1	<i>ez</i>	<i>min</i>
	2	<i>tu</i>	<i>te</i>
	3	<i>ew</i>	<i>wî</i> (m.) / <i>wê</i> (f.)
PL	1	<i>em</i>	<i>me</i>
	2	<i>hûn</i>	<i>we</i>
	3	<i>ew</i>	<i>wan</i>

In addition to the personal pronouns, Kurdish has an invariant reflexive pronoun *xwe* (dialectically also [xæ], [xo]). It is used for all persons and numbers, both as a personal pronoun and a possessor, when coreference with the subject of same clause is intended. In Standard K., the reciprocal pronoun is *hev* or *hevdû*, again used for all persons.

There are two demonstratives, *ew* ‘that’ and *ev* ‘this’, with considerable dialectal variation. The standard forms are illustrated in Table 10.

Table 10: Demonstratives in Standard Kurmanjî

	DIRECT		OBLIQUE	
	(all gender/numbers)	Sg. masc.	Sg. fem.	Plural
PROXIMATE	<i>ev</i>	<i>vî</i>	<i>vê</i>	<i>van</i>
DISTAL	<i>ew</i>	<i>wî</i>	<i>wê</i>	<i>wan</i>

In addition to the demonstratives, most dialects attach an additional suffixal or clitic marker to the nouns preceded by a demonstrative: in the singular *-e/-a* (regional variants, not gender related) and in the plural *-ene* / *-ana*. They only attach to the noun if it is the final element of the NP; if it is followed by a modifier in an ezafe construction, then the marker is not used.

- (3) (Antep-Adiyaman region, Haig and Öpentin 2018)

Va defter=na yē min=in
 DEM.PROX notebook=DEM.PL EZ.PL 1SG.OBL=COP.3PL
 ‘These notebooks are mine’

2.3. The structure of the NP

The basic structure of a NP in Kurmanjî is the following, where only N(oun) is obligatory:

Dem	Num	N-Ez	Poss	Ez	Adj
<i>ev</i>	<i>sé</i>	<i>kum-ên</i>	<i>min</i>	<i>yêñ</i>	<i>reš</i>
these	three	hat-EZ.PL	1SG.OBL	EZ.PL	black
'these three black hats of mine'					

Demonstratives have already been illustrated in Table 10. The numerals are given in Section 2.4. Possessors and descriptive adjectives follow the head, in that order if both are present, and obligatorily occur with an ezafe (cf. Schroeder 1999 for discussion of the NP in written Kurdish).

2.3.1. The ezafe construction

The ezafe construction is well-known from Persian, and is found, with certain variations, in all varieties of Kurdish. It may be either a bound morpheme (suffix or clitic, with as yet poorly researched dialectal variation in stress assignment) or an independent particle. Historically it goes back to an Old Iranian demonstrative/relativizer (cf. Haig 2011). In Standard K., traces of these pronominal origins are evident in the fact that the ezafe still inflects for gender and number, agreeing with its head noun in these categories, and in the fact that it occurs as the “free”, or “demonstrative” ezafe, discussed briefly below. We distinguish between simple linking ezafe constructions, and free or demonstrative ezafes.

A simple ezafe construction is found with any noun that is modified by an adjective, as in (4), or with a possessor, as in (5). An ezafe particle attaches to the head noun, and varies according to gender and number of that noun:

- (4) *bajar-ek-i mezin*
 town-INDF-EZ.M big
 ‘a big town’

- (5) *mal-a me*
 house-EZ.F 1PL.OBL
 ‘our house’

Nouns may also be modified by prepositional phrases, as in (6), or by relative clauses as in (7) and (8), in which case they also take the ezafe:

- (6) *dest-ên min yêñ ji ber serma-yê gelîš-i [...]*
 hand-EZ.PL 1SG.OBL EZ.PL from ADP cold-OBL.F split-PTCP
 ‘my hands which are cracked and split because of the cold’

- (7) *ev ri-ya ku tu di-d-i pēši-ya me*
DEM.PROX road-EZ.F REL 2SG IND-give.PRS-2SG front-EZ.F 1PL.OBL
‘This road that you make us take’
(Sarman 37)
- (8) *kur-ē wī yē li welat-ēn xeribī-yē*
son-EZ.M 3SG.OBL.M EZ.M in country-EZ.PL foreign.land-OBL.F
‘His son (who) is in foreign countries’

Possessors in ezafe constructions take the oblique case, adjectives in ezafe constructions remain uninflected. Compare the difference:

- gel-ê kurd* (people-EZ.M Kurdish) ‘the Kurdish people’
welat-ê kurd-an (country-EZ.M Kurd-OBL.PL) ‘the country of the Kurds’

If a possessor is coreferent with the same-clause subject, the reflexive pronoun *xwe* is obligatorily used in place of a personal pronoun:

- (9) *Ez li mal-a xwe me / mal-a *min*
1SG in house-EZ.F REFL COP.1SG / house-EZ.F 1SG.OBL
im
COP.1SG
‘I am at my house’ (lit. I am at self’s house / *I am at my house)

The forms of the ezafe in Standard K. are given in Table 11:

Table 11: Ezafe with the nouns *gund* ‘village’, *bra* ‘brother’, *jin* ‘woman’, *çira* ‘lamp’

Singular				Plural (masc. and fem.)	
masculine		feminine			
Def	Indef.	Def.	Indef.	Def.	Indef.
<i>gund-ê</i>	<i>gundek-î</i>	<i>jin-a</i>	<i>jinek-e</i>	<i>gund-ēn / -êt</i>	<i>gund-in-e</i>
<i>bra-yê</i>	<i>brayek-î</i>	<i>çira-ya</i>	<i>çirayek-e</i>	<i>jin-ēn / -êt</i>	<i>jin-in-e</i>
				<i>bra-yēn / -yêt</i>	<i>bra-n-e</i>
				<i>çira-yēn / -yêt</i>	<i>çira-n-e</i>

The plural forms with *-êt* are found mainly in the Behdinî (see Haig, this volume, chapter 3.3, §4). As mentioned above in connection with gender, the gender distinction in the ezafe following the indefinite marker *-ek* tends to weaken, with considerable uncertainty and inconsistency in the forms. In the spoken language, an ezafe may be omitted completely following nouns with indefinite *-ek*, and this can also be witnessed sporadically in the written language: *li ber derê kafeyek internetê* ‘in front of the door of an internet cafe’, with no ezafe following the initial head noun (Dirêj 2011: 21).

The presence of an ezafe on any noun suppresses the expression of oblique case on that noun. This is a very crucial fact of Kurmanjî syntax: it means that the ezafe itself is impervious to the external case of the entire NP. For example:

- (10) *Gund di nav [çiya-yêñ bilind] da ye*
 village ADP middle mountan-EZ.PL high ADP COP.3SG
 ‘The village lies between high mountains.’

- (11) *Li wê herêm-ê [çiya-yêñ bilind] he-ne*
 ADP DEM.OBL.M region-OBL.M mountan-EZ.PL high existent-COP.PL
 ‘There are high mountains there.’

The ezafe construction *çiyayêñ bilind* remains unchanged, regardless of the syntactic function of the entire NP. Thus in (10), we would expect an oblique case, because it is the complement of an adposition, while in (11) we would expect direct case, because it is the subject of an existential clause. But the presence of the ezafe *-yêñ* suppresses any overt expression of case on the noun. Overt case is, however, expressed when the ezafe construction is introduced with a demonstrative, which always expresses the case of the entire NP:⁴

- (12) *Gund di nav [wan çiya-yêñ bilind] da ye*
 village in middle those mountain-EZ.PL high ADP COP.3SG
 ‘The village is in between those high mountains.’

- (13) *[ew çiya-yêñ bilind] li ser sinor in*
 those mountain-EZ.PL high ADP ADP border COP.3PL
 ‘Those high mountains are on the border.’

Ezafe particles may also occur separated from their head noun. We refer to these as free, or demonstrative, ezafes. There are two possibilities. First, they may be used to add additional dependents to an existing simple ezafe construction. They still agree with the respective head nouns in number and gender:

- bra-yê min ê mezin* ‘my older brother’
mehîn-a boz a qenc ‘the good grey mare’
gund-êñ Qersê ên kevn ‘the old villages of Kars’

Second, they occur as anaphoric elements with the sense of ‘the one ...’. In such contexts, they are prosodically independent, rather than enclitic, and are preceded by a glide: *yê spî* ‘the white one (masc. sg.)’; *ya te* ‘your one (fem. sg.)’; *yêñ mezin* ‘the big ones (pl.)’.

⁴ Interestingly, in Şemzînan (and probably Badinan generally) this sometimes does not hold, and the demonstrative may actually remain in the direct case: *tu ew çiyayêt bilind dibîñi?* ‘Do you see those high mountains’, where the demonstrative is in direct case.

In Behdinî Kurdish, and dialects close to it such as Şemzinan (cf. §3.2.1), the use of ezafes as independent forms has entered the verbal domain, where they accompany certain kinds of predicates, in particular copular elements (cf. Haig 2011, and this volume, chapter 3.3, §4), as in (14).

- (14) ... *ez ya bêdeng im*
 I EZ.F silent COP.1SG
 ‘I am remaining silent’ (the speaker is a woman)⁵

Something similar may be found in the Elbistan dialect, discussed below in §3.2.2, though the origins of these forms are somewhat obscure.

2.4. Numerals

The main numbers, given in Standard Orthography (following Bedir Khan and Lescot 1970) are as follows:

1	<i>yek</i>	11	<i>yanzdeh, yazdeh</i>	30	<i>sî</i>
2	<i>didu, du</i>	12	<i>dwanzdeh</i>	40	<i>çel, cil</i>
3	<i>sisê, sê</i>	13	<i>sêzdeh</i>	50	<i>pêncî</i>
4	<i>çar</i>	14	<i>çardeh</i>	60	<i>şêst</i>
5	<i>pênc</i>	15	<i>panzdeh</i>	70	<i>heftê</i>
6	<i>şes</i>	16	<i>şanzdeh</i>	80	<i>heştê</i>
7	<i>heft [hæft]</i>	17	<i>hevdeh</i>	90	<i>nod, not</i>
8	<i>heşt [hæft]</i>	18	<i>hejdeh</i>	100	<i>sed</i>
9	<i>neh</i>	19	<i>nozdeh</i>	201	<i>du sed û yek</i>
10	<i>deh</i>	20	<i>bîst</i>	1000	<i>hezar</i>

The short forms of 2 and 3 are used when they are quantifiers in a NP: *sê zarok* ‘three children’. In the western parts of the Kurmanjî speech zone, the typically Indo-European opaque forms for 11 and 12 have disappeared, and all the numerals 11–19 have been regularized along the lines of “10-and-1”, “10-and-2” etc: *dehûyek, dehûdu, dehûsê* (cf. Haig 2006). This would appear to reflect contact influence from Armenian and Turkish, which lack opaque forms for 11 and 12, and instead have regularly-formed “10–1” etc.

⁵ From a short story *Hirmîka Xirş* by Mihemed Selim Siwarî, a writer from the Behdinî-speaking region in North Iraq, published in *Antolojiya çirokên kurmancêň başûr*, edited by Xelîl Duhokî (Avesta, 2011).

2.5. Adpositions

In Kurmanjî, I distinguish three components of the adpositional system, which can be combined in various ways: basic prepositions, locational nouns, and postpositional particles.

2.5.1. Basic prepositions

Standard Kurdish has three basic prepositions, and these are reasonably stable in most dialects (though Southeastern Kurmanjî differs in some respects, cf. §3.3). Each covers a broad and fairly abstract semantic space, with a spatial core: *ji* ‘from’, *bi* ‘by, through’, *li* ‘at’.

Recall from the discussion of phonology in §2.1 that no word can end with the short central vowel [i]. From this fact, it is evident that these prepositions do not constitute phonological words in Kurdish, and are probably best seen as proclitics. These three prepositions are also unique in that they fuse with a demonstrative to yield *jê* (*ji* + *wî/wê*), *pê* (*bi* + *wî/wê*) and *lê* (*li* + *wî/wê*) respectively. In some dialects, the compositional form *pê* has been reanalysed as a simple preposition with instrumental meaning, cf. (15) from Karakoçan dialect (field notes from Karakoçan):

- (15) *ister pê dest-an bi-xw-e ister pê kevčî*
want with hand-OBL.PL SUBJ-eat:PRES-IMPER.2SG want with spoon
‘Eat (it) with (your) hands, or with a spoon, as you please’⁶

Basic prepositions may occur alone, but are more common in combination with a postpositional particle, or with locational nouns. In addition to the three mentioned above, there is also one fixed circumposition, consisting of *di* ... *de* ‘inside’. In Standard K., the prepositional element *di*, unlike the three mentioned above, cannot occur by itself, but is always accompanied by the postpositional particle *de*.

In addition to the three “basic prepositions” just discussed, two other prepositions are found throughout the Kurmanjî region, though they differ from the three just mentioned in that they end in full vowels, and there is no fusion with the prepositional complement: (a) the preposition *bê* ‘without’; (b) the preposition *bo*. In most dialects of Turkey *bo* can be combined with *ji* to express benefactive meanings (*ji bo*). The dialects of the southeast use simply *bo*, which is also extended to cover recipient and goal meanings, where it generally replaces the combination *ji* ... *re/ra* of Standard K.

⁶ *ister* ... *ister* is a modified loan construction from Turkish, based on the Turkish verb *istemek* ‘want’

2.5.2. Locational nouns

A number of prepositions are evidently the result of the grammaticalization of nouns; they can be used both independently and in combination with the basic prepositions. The commonest are the following:

<i>nav</i> ‘inside’	<i>di nav ... de</i>	‘inside’
<i>ber</i> ‘front’	<i>li ber</i>	‘in front of’
	<i>ji ber</i>	‘because of’
<i>ser</i> ‘head’	<i>li ser</i>	‘on, upon, over’
<i>bin</i> ‘bottom’	<i>li bin, di bin ... de</i>	‘beneath, underneath’
<i>dû</i> ‘behind’	<i>li dû</i>	‘after’
<i>piş</i> ‘back’	<i>li piş</i>	‘behind’
<i>rex</i> ‘side’	<i>li rex</i>	‘next to, on the side’
<i>tenişt</i> ‘side’	<i>li tenişt</i>	‘by side’

Some of these locational nouns also occur in a kind of prepositional stranding construction, occurring without a complement and post-verbally:

- (16) *min nan da ber û hat-im*
 1SG.OBL bread give.PST.3SG front and come.PST-1SG
 ‘I put the food in front (of him) and came’

2.5.3. Postpositional particles

These particles are mostly combined with a preposition; they cliticize to the right-edge of the entire prepositional phrase, and provide additional meaning components to the phrase. However, the resultant meanings are not always transparent, for example (cf. Bedir Khan and Lescot 1970: 244–258 for a detailed list):

<i>bi ... re/ra</i>	‘accompaniment, together with’
<i>di ... re / ra</i>	‘through’
<i>ji ... re/ra</i>	‘for, to, benefactive/recipient’
<i>(ber) bi ... de</i>	‘towards’

In many dialects of central Anatolia (e. g. Dersim and Karakoçan) and the dialects of Armenia, the benefactive/recipient *ji ... re/ra* and the local *di ... de/da* are reduced to just the respective postpositional element: *mi(n)-ra* ‘to/for me’, *mal-da* ‘at home’.

Spatial meanings are also conveyed through directional adverbs, such as: *jêr* ‘down’, *jor* ‘up’, *xwar* ‘down (on the ground)’. Another important element is the particle *-de/da* following NPs expressing directionals, when they occur after the predicate.

- (17) *hinek av-ê bi-xi-yê da ji_bo_ku*
 a.little water-OBL.F SUBJ-drop.PRS-3SG.OBL ADP so.that
ne-şewit-e
 NEG-burn.PRS-3SG
 ‘Put a little more water in it so that it does not burn.’

2.6. Verbal morphology

Verbs exhibit the typical western Iranian characteristic of having two stems, a present and a past stem, but the formation of one from the other is not fully predictable. Certain regularities can be identified, though we will not attempt a classification here. Table 12 provides a list of frequent verbs:

Table 12: Frequent verbs in past and present stems

Infinitive	Past Stem	Present Stem	Meaning
bûn	<i>bû-</i>	- <i>b-</i>	be
birin	<i>bir-</i>	- <i>b-</i>	take
hatin	<i>hat-</i>	- <i>(h)ê-/wer-</i>	come
hêlan/hîstîn	<i>hişt-</i>	- <i>hêl-</i>	leave
bihîstîn	<i>bihîst-</i>	- <i>bihîz-</i>	hear
girtin	<i>girt-</i>	- <i>gir-</i>	grasp, hold
gotin	<i>got-</i>	- <i>béj-</i>	say
kuştin	<i>kuşt-</i>	- <i>kuj-</i>	kill
rûniştin	<i>rûnişt-</i>	- <i>rûn-</i>	sit
kirin	<i>kir-</i>	- <i>k-</i>	do, make
çûn	<i>çû-</i>	- <i>ç-/her-</i>	go
jîn/jîyan	<i>jî-/jiya-</i>	- <i>jî-</i>	live
ketin	<i>ket-</i>	- <i>kev-</i>	fall
xwarin	<i>xwar-</i>	- <i>xw-</i>	eat
xwastin	<i>xwast-</i>	- <i>xwaz-</i>	want, request
avêtin	<i>avêt-</i>	- <i>avêj-</i>	throw
dîtin	<i>dît-</i>	- <i>bîn-</i>	see
dan	<i>da-</i>	- <i>d-</i>	give
mirin	<i>mir-</i>	- <i>mir-</i>	die
zanîn	<i>zanî-</i>	- <i>zan-</i>	know
girîn	<i>girî-</i>	- <i>girî-</i>	cry, weep
ajotin	<i>ajot-</i>	- <i>ajo-</i>	drive
barîn	<i>barî-</i>	- <i>bar-</i>	rain
xistin	<i>xist-</i>	- <i>x-/xîn-</i>	strike, knock
xwandin	<i>xwand-</i>	- <i>xwîn-</i>	read, study

Verbs are quite a small, closed word class in Kurmanjî (probably no more than 150 simplex verbs in regular usage in most dialects). The only moderately productive derivational process for creating new verbs is a causative suffix, *-and*, used for

deriving transitive verbs from intransitive present stems: *gerîn* ‘walk, stroll’ → *gerandin* ‘lead’, *nivistin* ‘sleep’ → *nivandin* ‘put to sleep’. New verb meanings are normally created using light verb constructions usually based on *kirin*, *bûn*, *dan* (see below). Additional verbs may also be derived through the lexicalization of verb plus a dummy prepositional complement, for example *jê birin* ‘win’, literally ‘take from him/her’, also ‘erase’; *lê xistin* ‘beat (a person), lit. ‘strike on him/her’; *lê hatin* ‘become’. Incorporation of a pre-verbal element may also yield a new verb (see below).

2.6.1. Person marking suffixes

Finite verbs take agreement suffixes, indexing the verb for person and number of a single core argument: the intransitive subject in all tenses, the transitive subject in present tenses, and the transitive object with past tenses. Table 13 shows the two sets of person agreement suffixes, one used with forms based on the present stem, the other for forms based on the past stem. Subjunctive forms based on the past stem have distinct composite endings with considerable cross-dialect variation, beyond the scope of this section (see the dialect sketches below for some discussion). Non-verbal predicates take a (slightly) different set of clitic copular endings given in Table 15.

Table 13: Person agreement suffixes

PERSON	STEM	
	PRESENT	PAST
1SG	- <i>(i)m</i>	- <i>(i)m</i>
2SG	- <i>i</i> / - <i>e</i> (imperative)	- <i>(y)i</i>
3SG	- <i>e</i>	-Ø
1,2,3PL	- <i>(i)n</i>	- <i>(i)n</i>

TAM and negation prefixes

Verb forms based on the present tense (simple present, subjunctive, imperative, future) obligatorily take a single prefix, either the neutral simple present prefix *di-* (glossed INDICATIVE), or the subjunctive prefix *bi-*, or a negation prefix *na-* (replaces the indicate)⁷ or *ne-* (replaces the subjunctive). In imperative forms, and with preverbal elements combined with light verbs, a prefix may be lacking. In

⁷ Two verbs negate the present stem with *ni-*, *zanîn* ‘know’ and *karîn* ‘be able’: *nizanim* / *nikarim* ‘I don’t know / I can not’. The verb *şiyân* ‘be able’, used in Behdinî and Şemzînan dialects, negates the present stem with *ne*: *neşêm* ‘I can not’.

the Behdini dialects of North Iraq, the subjunctive present forms used to make the future tense regularly drop the prefix *bi-*.

Verb forms based on the past stem, however, are not necessarily prefixed. The simple past tense is basically thus the past stem of the verb plus the appropriate person agreement markers. In the past tenses, an aspect distinction between progressive (or imperfective) and simple past is available, signalled by the prefix *di-*.

Negation of both simple and progressive pasts⁸ is through addition of the prefix *ne-*: *ne-hatin* ‘they didn’t come’, or *ne-di-hatin* ‘they weren’t coming’. In past tenses, agreement patterns vary according to the transitivity of the verbs (cf. §3.4–3.5 below). Sample paradigms are given in Table 14.

Table 14: Sample verb conjugations

	simple present (trans.) <i>gotin</i> ‘say’	simple past (intrans.) <i>hatin</i> ‘come’	simple past (trans.) <i>xwarin</i> ‘eat’	past progres- sive (intrans.) <i>hatin</i> ‘come’	past progressive (trans.) <i>xwarin</i> ‘eat’
1SG	<i>ez dibêjim</i>	<i>ez hatim</i>	<i>min xwar</i> ‘I ate (sth.)’	<i>ez dihatim</i>	<i>min dixwar</i>
2SG	<i>tu dibêjî</i>	<i>tu hatî</i>	<i>te xwar</i>	<i>tu dihatî</i>	<i>te dixwar</i>
3SG	<i>ew dibêje</i>	<i>ew hat</i>	<i>wî</i> (masc.) <i>xwar</i> <i>wê</i> (fem.) <i>xwar</i>	<i>ew dihat</i>	<i>wî</i> (masc.) <i>dixwar</i> <i>wê</i> (fem.) <i>dixwar</i>
1PL	<i>em dibêjin</i>	<i>em hatin</i>	<i>me xwar</i>	<i>em dihatin</i>	<i>me dixwar</i>
2PL	<i>hûn dibêjin</i>	<i>hûn hatin</i>	<i>we xwar</i>	<i>hûn dihatin</i>	<i>we dixwar</i>
3PL	<i>ew dibêjin</i>	<i>ew hatin</i>	<i>wan xwar</i>	<i>ew dihatin</i>	<i>wan dixwar</i>

Non-verbal predicates in the present indicative receive a set of clitic person markers or copula forms, see Table 15.

Table 15: Copular person clitics with non-verbal predicates

	Following a consonant: <i>kurd</i> ‘Kurdish’	Following a vowel: <i>birçî</i> ‘hungry’
1SG	<i>ez kurd-im</i> ‘I am Kurdish’	<i>ez birçî-me</i> ‘I am hungry’
2SG	<i>tu kurd-i</i>	<i>tu birçî-yî</i> (reduced to [i:j])
3SG	<i>ew kurd-e</i>	<i>ew birçî-ye</i>
1PL	<i>em kurd-in</i>	<i>em birçî-ne</i>
2PL	<i>hûn kurd-in</i>	<i>hûn birçî-ne</i>
3PL	<i>ew kurd-in</i>	<i>ew birçî-ne</i>

⁸ Negation of progressive past in Western Kurmanji (Adiyaman-Urfa) is *na-*, as in *na-dekir-in* ‘they were not doing it’.

In Western Kurmanjî, special constructions are found with non-verbal predicates (cf. §3.4.2.2). For non-verbal predicates in the past tenses, or in subjunctive mood, the appropriate form of *bûn* ‘be’ is required.

2.6.2. The verbs *çûn* ‘go’ and *hatin* ‘come’

These two verbs have suppletive stems, with regional variation in the choice and forms of the stems.

Table 16: The verbs *çûn* ‘go’ and *hatin* ‘come’

	<i>çûn</i> ‘go’	<i>hatin</i> ‘come’
Present	- <i>çî-</i> / - <i>her-</i> (imperative)	-(<i>h</i>) <i>ê-</i> / - <i>wer-</i> (imperative)
Past	<i>çû</i>	<i>hat</i>

The Standard K. indicative singular 1SG form of ‘come’ is *têm*, resulting from a contraction of **di-hê-m* with the typical devoicing of the *d*- in such contexts (in some dialects the trace of the stem-initial *h*- can still be heard), while the negation is *nayêm*. In the western dialects of Kurmanjî (WK, see §3.4.2 below), the imperative stem -*her-* of ‘go’ is also used in the indicative, so for example in Elbistan, Dersim, Erzurum, and Elazığ, it is used for all forms of the present stem. Thus first person indicative present in these dialects is *terim* (<**di-her-im*) ‘I go / am going’, while negative indicative present is *narim* (<**na-her-im*). In other dialects, the imperative form is also used to cover subjunctive meanings in the present. The imperative stem of *hatin* is also often used in place of the regular subjunctive (which is *bêm* in the first singular).

2.6.3. Mood

With the present stem, there is a simple distinction between indicative verb forms, marked with *di-*, and subjunctive verb forms, prefixed with *bi-* or zero in some dialects.

The subjunctive of the present stem (cf. 18–21) has a wide range of functions, including clauses with irrealis sense (wishes or orders), and subordinate clauses expressing possible or intended outcomes. It is obligatory in the complements of modal predicates such as ‘want’, ‘be able’, ‘be obliged/must’. Some examples follow (from Bedir Khan and Lescot 1970: 317–321):

- (18) *Kafîr jî b-e*
 unbeliever ADD be.PRS.SUBJ-3SG
 ‘even if (he) is an unbeliever’

- (19) *Ez di-tirs-im şermisarî û belengazi*
 1SG AFF-fear.PRS-1SG disgrace and misery
par-a me b-e
 fate-EZ.F 1PL.OBL be.PRS.SUBJ-3SG
 ‘I am afraid that disgrace and misery will be our fate’
- (20) *heke birçî ne, bila bi-xw-in*
 if hungry COP.PL MOD.PRT SUBJ-eat.PRS-PL
 ‘If they are hungry, they should eat’
- (21) *heke pirs-a wan he-b-e,*
 if question-EZ.F 3PL.OBL existent-be.PRS.SUBJ-3SG
bila vê gave bêj-in
 MOD.PRT this time-OBL say.PRS.SUBJ-PL
 ‘If they have a question, they should say so at the time’

Other subordinate clauses may (as in (22) and (23)) or may not be in the subjunctive (as in 24–25), depending on the degree of certainty of the proposition expressed.

- (22) *Dît ko gotin-a wî rast e*
 see.PST(3SG) COMPL word-EZ.F 3SG.M.OBL right COP.3SG
 ‘He saw that what he said was right’
- (23) *De bêje, te çawan kir*
 MOD.PRT say.PRS.IMPER 2SG.OBL how do.PST.3SG
 ‘go on, say how you did it’
- (24) *Ni-zanîbû⁹ ko çawan bê pere*
 NEG-know.PST(3SG) COMPL how without money
ve-ger-e mal
 PRV-return.PRS.SUBJ-3SG home
 ‘He didn’t know how he would return home without any money’
- (25) *Bi-xwîn-in da ko ho zana bi-b-in.*
 SUBJ-read.PRS-PL so that thus knowledgeable SUBJ-be.PRS-PL
 ‘study, so that you may become knowledgeable’

2.6.4. The future tense

The future tense is expressed via a particle (*d)*ê or *w*ê, combined with a subjunctive form of the verb; see (26). It is most frequent in clause-second position, often (but not always) preceded by the subject as in (27), though it is also possible clause-

⁹ The verb *zanîn* ‘know’ usually takes this form for the negated simple past.

initially; see (28) below. The main verb is in the present subjunctive. High-frequency verbs such as *gotin* ‘say’ generally drop the *bi-* prefix in the future tense, as does the verb *bûn* ‘be’ (exs. from Haig and Öpentin 2018).

- (26) *ji_bo Xwedē sal-ek-ē ez=ē řožt̄ bi-gir-im*
for God year-INDF-OBL 1SG=FUT fasting SUBJ-keep.PRS-1SG
'For the sake of God I will fast during one year.'
- (27) *ew dē mesel-ē ji te re bēz-e*
3SG FUT issue-OBL.F from 2SG.OBL POSTP say.SUBJ.PRS-3SG
'He will tell you the issue.'
- (28) *wē čawa heval-ēn nexweš derbas bi-bi-n*
FUT how friend-EZ.PL ill PRV.pass SUBJ-be.PRS-3PL
'How will the wounded friends pass?'
- (29) *ewro ne, dē sibe či-m*
today no, FUT tomorrow come.SUBJ.PRS-1SG
'not today, (but) tomorrow I will come'

In Standard K. and contemporary written Kurdish, the future auxiliary can be retained in negative sentences, in which case the negation marker is the subjunctive negation ‘*ne-*’. In Behdinî, there is no dedicated negative future; the negative indicative is used:

- Standard K.: *Ez ē sibe bi wan re ne-či-m.*
Behdinî: *Ez sibe digel wan na-či-m* ‘I won’t go with them tomorrow’

2.6.5. The directional *-e* particle on verbs

A large class of verbs expressing motion (‘go’, ‘come’) or directed action (‘give’, ‘speak’, ‘send’) frequently takes the so-called ‘directional particle’ *-e* (in most dialects [-æ]) after the person marker on the verb. MacKenzie (1961a: 197–198) analyses it as a reduced form of a directional preposition, which has cliticized to the preceding verb. It is assimilated to the final vowel of verb forms ending in one of the full vowels (e. g. *čû* ‘went’). For verbs such as *hatin* ‘come’, the use of the directional is almost obligatory (claimed, for example, for Tur Abdin dialect in Turgut 2012). An example from a traditional text is the following (Bedir Khan and Lescot 1970: 352):

- (30) *Se û zarok-ên gund li gur hat-in-e*
 dog and child-EZ.P village at wolf come.PST-3PL-DRCT
hev ...
 together ...
 ‘The dogs and the children of the village gathered together around the wolf ...’

The precise conditions determining its realization remain, however, not fully understood. It needs to be distinguished from the reduced form of a third person singular addressee or recipient, which likewise cliticizes to the verb: *got=ê* ‘said to him/her’, *didin=ê* ‘give to him/her’. If such a clitic goal is present, the directional particle cannot be realized.

2.6.6. Light verb constructions

Like most Iranian languages, Kurdish makes extensive use of complex predicates consisting of a so-called ‘light verb’ plus some non-verbal element. The most commonly used light verbs in Kurmanjî are *kirin* ‘do, make’, *bûn* ‘be, become’ and *dan* ‘give’. The following list is a small selection of widely-used light verb constructions involving a nominal non-verb element (Haig 2002: 22–23):

<i>ava kirin</i>	‘build, establish’	<i>xwedi kirin</i>	‘bring up, raise’
<i>bang kirin</i>	‘call’	<i>bar kirin</i>	‘load, move (house)’
<i>alî(karî) kirin</i>	‘help’	<i>gazî kirin</i>	‘call’
<i>bawer kirin</i>	‘believe’	<i>hez kirin</i>	‘like, love’
<i>fa(h)m/fâm kirin</i>	‘understand’	<i>dest pê kirin</i>	‘start, begin’
<i>şerm kirin</i>	‘be ashamed’	<i>nîşan dan</i>	‘show’
<i>ji bîr kirin</i>	‘forget’	<i>dest avétin</i>	‘reach for, begin’
<i>guhdarî kirin</i>	‘listen’	<i>dev jê berdan</i>	‘leave alone, cease doing’

2.7. Syntax of the simple clause

2.7.1. Word order

The word order in pragmatically neutral clauses is SOVG, where “G” stands for “Goal”, here a cover term for spatial goals of verbs of movement, recipients of verbs of transfer, and addressees of verbs of speech. However, word order is not rigidly fixed; direct objects may be fronted for pragmatic purposes, for example. The position and means of marking of goal arguments (in the broad sense just defined) also varies; in those dialects which make extensive use of the circumposition *ji* ... *ra* (cf. §3.5) for recipients and benefactives, they precede the verb, yielding SGOV. For recipients with *dan* ‘give’, however, all dialects usually place

the recipient argument immediately after the verb, in the oblique case but with no adposition. The southeastern dialects make more extensive use of the post-predicate position, which is almost always coupled with the presence of the directional particle on the verb (§2.6.5). These dialects also make use of a preposition *bo* with some post-predicate recipients and benefactives (see Haig 2015 on post-predicate goals), in which case no directional particle occurs on the verb.

2.7.2. Alignment and related issues

Kurmanjî has an ergative construction, used with the past tenses of transitive verbs. Otherwise, the syntax is accusative throughout. The ergative construction associated with past transitive verb forms has attracted a fair bit of attention in recent years (Bynon 1979; Dorleijn 1996; Matras 1997; Haig 1998; Turgut 2012; Haig 2008 for summary discussion), and we will only point out some of the more salient facts here, and some points of variation across the dialects.

In the ergative construction, the transitive subject takes the Oblique case, while the direct object is in the Direct case. The verb agrees with the direct object. However, the order of subject and object remains unchanged. Similarly, the subject, despite its Oblique case, still controls coreference with reflexive *xwe*. Throughout Northern Kurdish, it is subjects only which control reflexive *xwe*, with exceptions only regularly found in WK (§3.5). Thus the relation of subjecthood in Northern Kurdish is quite robust, and largely independent of surface case (cf. Haig 1998 for discussion). Typical examples (from Thackston 2006: 49) are given in (31–33):

- (31) *Jinik-ek-ê çay-a me anî*
woman-INDF-OBL.F tea-EZ.F 1PL.OBL bring.PST.3SG
'A woman brought our tea.'
- (32) *Wî mirov-î çay anî*
DEM.M.OBL man-OBL.M tea bring.PST.3SG
'That man brought tea.'
- (33) *Gundi-yan tişt-ek ne-got*
villager-PL.OBL thing-INDF NEG-say.PST.3SG
'The villagers didn't say anything.'

There are two main areas where the morpho-syntax of Kurdish diverges from what would be expected from the rules of ergativity as just outlined, namely the agreement on the verb, and the case of the direct object. With regard to verb agreement, when the transitive subject is plural, and not expressed overtly in the clause, there is a strong tendency to add plural agreement to the verb, even when the object is singular. This usage is followed in all varieties of Kurdish, spoken or written, when the clause with the transitive verb is preceded by an intransitive clause with the same subject. Characteristically in (34) below, which represents the written

language (the poem *Ji Biçukan re*, by Cegerxwîn), an intransitive clause precedes the past transitive verb *gotin* ‘say’:

- (34) *Herdu çû-n-e cem rovî Doz-a xwe jê_ra
the.two go.PST-PL-DRCT to fox case-EZ.F SELF to.him
got-in
say.PST-PL*
 ‘The two of them **went** to the fox (and) **explained(PL)** their case to him.’

A second tendency, found in the dialects of Central Anatolia to the west, is to put the direct object of a past transitive verb into the Oblique case, rather than the expected Direct case, leading to a double-oblique construction (with both subject and object in the Oblique). In the dialect of Muş, this tendency can be regularly observed:

- (35) (story told by speaker from Muş)
*ez zarok bû-m-e, biçûk bû-m-e, min
1SG child be.PST-1SG-PERF small be.PST-1SG-PERF 1SG.OBL
girt-in-e ...
take.PST-3PL-PERF*
 ‘I was a child, I was young (they) took **me** ...’

See Haig and Öpentin (2018), and Haig (2017: 477–479) for discussion and references on deviations to canonical ergativity.

2.7.3. Non-canonical subjects

In the dialects of the south and the east (e. g. Şemzînan and Behdinî), certain predicates take a subject in the Oblique, regardless of tense. Such constructions resemble superficially the ergative construction, but should not be confused with it, because (a) they are not conditioned by the tense of the verb; (b) the predicates concerned can be intransitive. Typically such non-canonical subjects occur with certain predicates of physical sensations, for example *min(OBL) sar e* ‘I am cold’. However, not all such predicates have non-canonical subjects, cf. *ez(DIR) birčî me* ‘I am hungry’. The verb *viyan*, expressing necessity/desire, also takes an oblique ‘wanter’:

- (36) *min d-vê-t b-çi-m
1SG.OBL IND-be.necessary.PRS-3SG SUBJ-go.PRS-1SG*
 ‘I want to go’

Finally, in expressions of possession the possessor is often in the oblique:

- (37) *min trimbêl nîne
1SG.OBL car not.existent.3SG*
 ‘I do not have a car.’

In most other dialects, these constructions are not used. Instead, canonical subjects in the Direct case are used, or, in the case of possession, the possessor is the modifier in an ezafe construction. A remnant of this construction may be found in many dialects in the expression *çav ketin* ‘eye fall’, i. e. ‘catch sight of’, where the ‘possessor’ of *çav* occurs clause-initially, rather than via an ezafe-construction. The following example from Ritter’s Midyat texts (transcription adapted) is fairly typical:

- (38) *waxtē ku sofî çav pê ket* ...
 time-EZ.M COMPL Sufi eye with.him fall.PST.3SG
 ‘When the Sufi caught sight of him ...’

3. Dialectal variation in phonology and morphosyntax

This section briefly summarizes aspects of regional variation, largely based on Haig and Öpentin (2018). For more detailed discussion of lexical and phonological variation, see Öpentin and Haig (2014), which is based on a comparative list of lexical items. The two studies just mentioned yield a broad division of Kurmanjî into three main dialect groups: Southeastern Kurmanjî (SEK), Southern Kurmanjî (SK), and Western Kurmanjî (WK). Their approximate respective locations are indicated in Figure 3 below (see §4 for details).

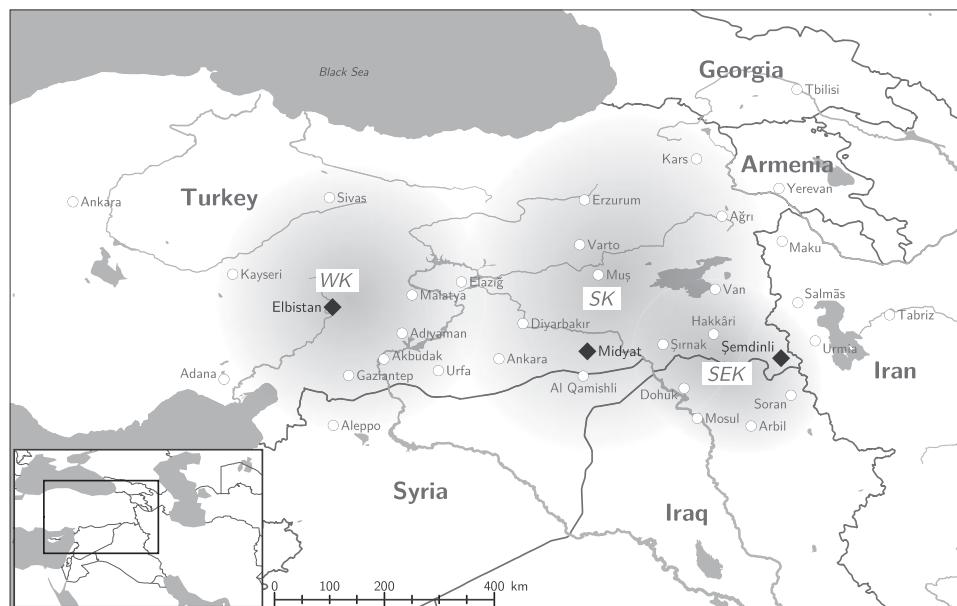


Figure 3: Approx. locations of three main Kurmanjî dialect zones

The situation is best captured in terms of a northwest-to-southeast dialect continuum, with the two endpoints Southeastern Kurmanjî and Western Kurmanjî as the most clearly differentiated dialects. In the intermediate region, provisionally termed here ‘Southern Kurmanjî’ (SK), dialect boundaries are blurred, and features of both WK and SEK are found to varying degrees. Probably the most divergent dialect is SEK, which includes Behdinî of North Iraq (see Haig, this volume, chapter 3.3, §4).

An initial impression of dialectal diversity can be gained by considering the distribution of the lexical item used to express the English activity verb ‘speak’. In Kurmanjî, at least eight distinct lexical items (including complex predicates) are used to express the English verb ‘speak’. They are provided in Table 17:

Table 17: Lexical expressions for SPEAK (numbers in brackets refers to numbering in Figure 4)

<i>deyn/deng kirin</i> (1)
<i>staxilîn</i> (2)
<i>xeber dan</i> (3)
<i>axaftin</i> (4)
<i>peyivîn</i> (5)
<i>qise kirin</i> (6)
<i>qezi kirin</i> (7)
<i>sor kirin</i> (8)

The distribution of these eight variants is largely geographically determined, and is graphically represented in Figure 4. Each triangle represents the location of a speaker, while the numbers refer to the variant of ‘speak’ which she used in response to a translation task (see Table 17 for the actual variants). Most of the data were extracted from the Database of Kurdish dialects (Matras et al. 2016); see Haig and Bulut (2017) for details of the methodology and the map.

It is evident that SEK fairly consistently uses variant 4 *axaftin*, while WK uses fairly consistently 1 *deyn kirin*. Elsewhere areal trends are clearly visible, but we also find, for example, that variant 3 *xeber dan* is widely used throughout the central region.

An area of morphosyntax where the dialectal divisions mentioned in Figure 3 are also relevant is adpositions, and word order. A particularly clear example concerns the position, and adpositional marking, of the Addressee argument of the verb ‘tell, say’, expressed through the lexeme *gotin* (with some phonological variants) in all dialects of Kurmanjî. Three main constructions are associated with this verb, and are illustrated in examples (39–41). In (39), typical for SEK, the Addressee is post-verbal, and the verb carries the so-called directional particle (see §2.6.5), the attenuated remnants of an earlier preposition. In (40), typical for Standard K. and most of the core of the Kurmanjî speaking zone in Anatolia, the Addressee

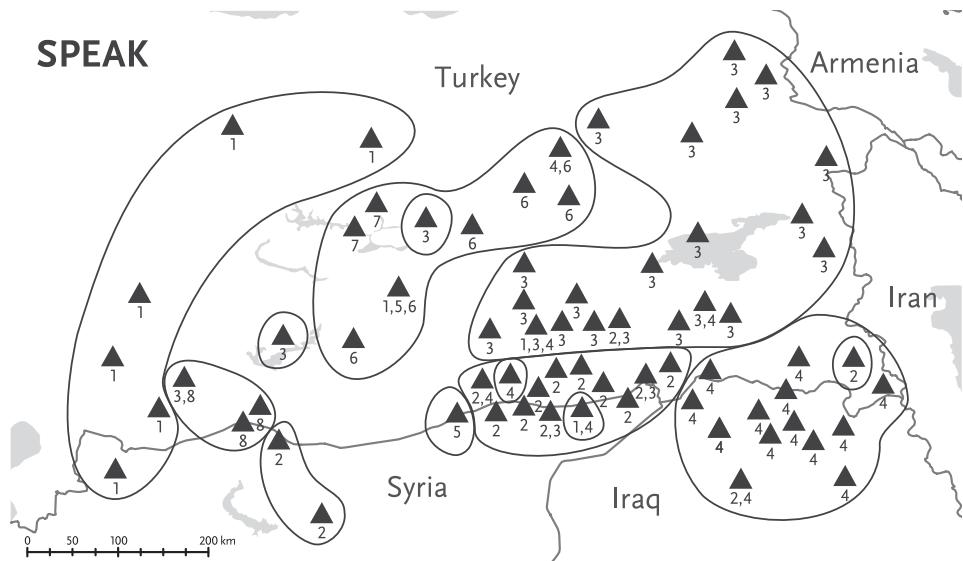


Figure 4: Distribution of lexical variants for ‘speak’ (from Haig and Bulut, 2017)

is pre-verbal, and flagged through a circumposition, *ji* ... =*ra*. Finally, in (41), typical of WK and the northern peripheries of Anatolia, we find the Addressee flagged solely through the postpositional clitic =*ra*, clearly a reduced form of the circumpositional variant illustrated in (40), via loss of the preposition.

The three examples each represent translations of the sentence ‘She/that woman told me’, and were extracted from the data for their respective locations from the Database of Kurdish Dialects (Matras et al. 2016). Note that *gotin* is a transitive verb in Kurmanji, hence the subjects are in the Oblique case in these past-tense clauses. Figure 5 is a map compiled by the present author to indicate the areal distribution of these three variants, with each point indicating the location of a speaker from the Database of Kurdish Dialects (Matras et al. 2016).

- (39) Şemzinan (SEK)

<i>ewê</i>	<i>got=e</i>	<i>min</i>	
3SG.OBL.F	tell.PST.3SG=DRCT	1SG.OBL	
'She told me'			

(40) Bingöl (SK)

<i>wê</i>	<i>jin-ê</i>	<i>ji</i>	<i>mi=ra</i>	<i>gotibû</i>
3SG.OBL.F	woman-OBL.F	ADP	1SG.OBL=ADP	tell.PPRF.3SG
'That woman told me'				

(41) Elbistan (WK)

wē mi=ra go
 3SG.OBL.F 1SG.OBL=ADP tell.PST.3SG
 ‘She told me’

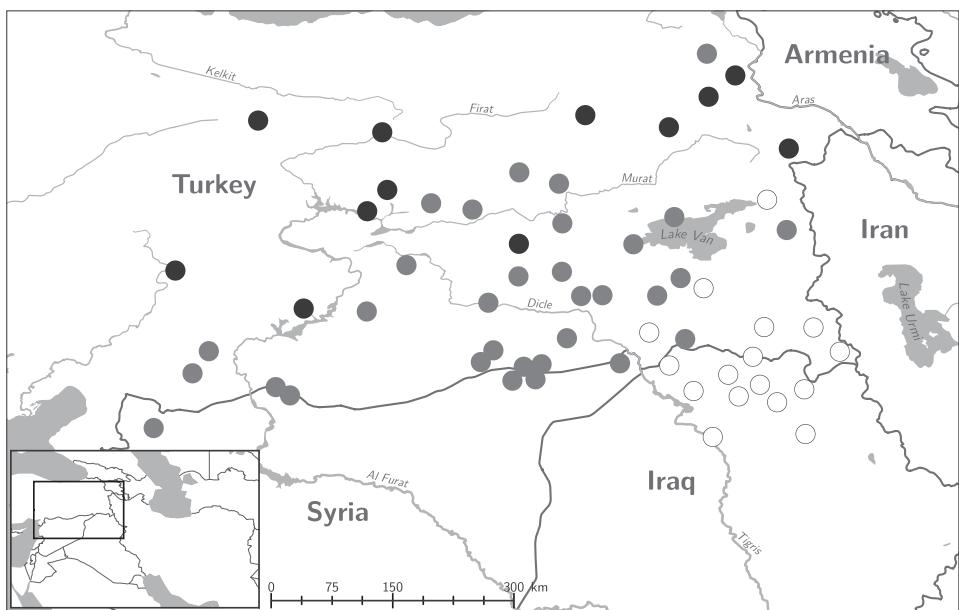


Figure 5: Distribution of construction types with the verb *gotin* in Kurmanji
 Key: white=post-verbal, with directional particle (39); grey=pre-verbal, with circumposition (40); black=pre-verbal, with postposition (41).

Having briefly illustrated the main dialectal divisions within Kurmanji, I will turn to some more specific features of phonology and morphosyntax, focussing on features where SEK and WK show the most divergent features with respect to Standard K., and what has been provisionally termed Southern Kurmanji above. Again, it must be stressed that Southern Kurmanji is not a well-defined entity, but essentially covers those areas that are not captured by SEK and WK. Within SK, there is of course a great deal of finer-grained dialectal differentiation, but it is beyond the scope of this chapter to cover it.

3.1. Variation in Phonology

3.1.1. Phonological variation in SEK

In SEK of the Şemzinan dialect, there is a process of final-vowel centralization: A high front unrounded vowel [i:] is mostly centralized into an [ɪ] in word-final position. The process affects certain inflectional morphemes, for example Standard Kurmanjî *wî mirov-i* ‘that man-OBL.M’ is *wi mirow-i* in Şemzinan, Standard K. *naç-i* ‘NEG.go.PRS-2SG’ is *naç-i* in Şemzinan. It also affects some lexical items, e. g. *tiji* as opposed to Standard K. *tijî*. However, otherwise a lexical final long [i:] is preserved, as in *spî* ‘white’, *tarî* ‘dark’, or *kari* ‘a sort of plant’. Note that the masculine ezafe and the oblique case following the indefiniteness suffix are not subject to centralization, as seen in the following examples:¹⁰

<i>mirov-ek-î baş</i>	(man-INDF-EZ.M good)	‘a good man’;
<i>xanî-yê mirov-ek-î</i>	(house-EZ.M man-INDF-OBL.M)	‘a man’s house’

This feature is salient in the eastern half of the SEK dialect zone, but not found in the western section such as in Dohuk or Amêdî (see Haig, this volume, chapter 3.3, §4).

Another feature of SEK, particularly of northern Iraq (Haig, this volume, chapter 3.3, §4.1.2), is the fronting of [u:] towards [y:]. In Behdinî of North Iraq and in the southeastern dialects of Kurmanjî in Turkey, the fronting process is accompanied by de-rounding, leading to [i:] in a number of lexical items, e. g. [xæsi:] ‘mother-in-law’, or [di:r] ‘far’ (Zakho, North Iraq), as opposed to Standard Kurmanjî [xæsu:] and [du:r]. Examples of fronting of [u:] to [y:], transcribed here as <ü>, are given in Table 18 (examples from Şemzînan dialect):

Table 18: Vowel fronting in Southeastern Kurmanjî (SEK)

Şemz.	Standard K.	Gloss
<i>stûr</i>	<i>stûr</i>	‘thick’
<i>mû</i>	<i>mû</i>	‘hair’
<i>tû</i>	<i>tû</i>	‘mulberry’
<i>bičûk</i>	<i>bičûk</i>	‘child’
<i>bük</i>	<i>bük</i>	‘bride’

Bilabialization of the voiced labio-dental fricative: Standard K. [v] is systematically seen as an approximant [w] in Şemzînan, similar to much of Central Kurdish (Haig, this volume, chapter 3.3, §3.1), so the the contrast between [v] and [w]

¹⁰ The final [i:] of participles is centralized when the participle is used predicatively (discussed in §3.2.1), as in: *mala wan a soti* ‘their house has burnt down’. Otherwise, participles retain the final long vowel.

is neutralized. For example, Standard K. *av*, *sev*, *çav* ‘water, night, eye’ are pronounced *aw*, *sew*, *çaw* in Şemzînan. The process can also be observed in loanwords, such as *vazo* ‘vase’ (from Turkish), which is *wazo* in Şemzînan. The few occurrences of [v], as in *vize viz* ‘swirling of flies’ and *bive ye* ‘it’s dangerous’ in child language, are onomatopoeic words and the speakers are usually not systematic in their pronunciation. In the Behdinî dialect of Dohuk, however, lenition of Standard K. [v] is not evident, and in fact the [v] in syllable-final position tends to be devoiced to [f]. Thus in these dialects, an opposition between [v] and [w] is retained (Haig, this volume, chapter 3.3, §4).

3.1.2. Phonological variation in WK

The phonology of WK diverges from that of Standard K. in several respects. The most striking is the backing and rounding of Standard K. [a:] to WK [ɔ:], shown in Table 19.

Table 19: Backing and rounding of /a/ in Western Kurmanjî (WK)

Orthog.	Standard K.	WK	Gloss
<i>av</i>	[a:v]	[ɔ:v]	‘water’
<i>hatin</i>	[ha:tin]	[ho:tin]	‘to come’
<i>da</i>	[da:]	[dɔ:]	‘s/he/it gave’

Standard K. [ɛ] or [æ] is also regularly retracted to a low central unrounded vowel [œ] in WK. Thus Standard K. *dest* ‘hand’ and *dev* ‘mouth’ ([dæst], [dæv]) become [dəst] and [dəv] respectively.

Turning to the consonants, the Standard K. [b] is lenited via [β] into an approximant [w] in intervocalic, and in some cases, in word-initial and word-final positions. Note that the phenomenon is restricted to intervocalic position in other dialects (such as northern part of Kurmanjî speech zone). The process regularly affects an initial [b-] of verbal stems, when they are preceded by a tense, aspect, mood, or negation prefix.

Table 20: Lenition of pre-vocalic /b/ in Western Kurmanjî (WK)

Orthog.	Standard K.	WK	Gloss
<i>hebek</i>	[hæbæk]	[hæwæk]	‘one unit’
<i>seba</i>	[sæba:]	[səwa:]	‘because of’
<i>bîne</i>	[bi:næ]	[wi:nœ]	‘Bring (it)!’
<i>bibîne</i>	[bibî:næ]	[biwi:ni]	‘(If s/he) sees (it)’
<i>nebêze</i>	[næbe:zæ]	[mæwe:]	‘Do not say!’
<i>kîtêb</i>	[kʰite:b]	[kʰite:w]	‘book’

An epenthetic vowel [i] (see §2.1) in a number of Standard K. function words and inflectional morphemes is regularly a full vowel [æ] in WK (similar to Sorani/Central Kurdish to which, geographically, WK is the most distant region). For example, the indicative present suffix *di-* generally loses its vowel, and is reduced to [d-], or [t-] in e. g. SEK (the entire morpheme is generally absent in the Mardin region of Southern Kurmanjî), just as the epenthetic vowel of the basic prepositions is likewise often lost in SEK. In WK, on the other hand, we find the indicative prefix [dæ-], as in *dæ-kim* ‘I do’, or the preposition *la* for SEK [l(i)] ‘at’.

Notice that a pharyngeal [χ] is altogether not attested in the data of the Elbistan variety of WK dialect. That is, the few words which are most prone to the development of pharyngeals in Kurdish dialects, such as Standard K. *mar* ‘snake’, *tehl* ‘bitter’, *čav/čehv* ‘eye’, do not contain a pharyngeal phoneme.

3.2. Variation in the Ezafe construction

3.2.1. Ezafe in SEK

With definite nouns, the singular ezafe forms are the same as in Standard K., but there are some differences in the plural ezafe and elsewhere, summarized below, see also Haig (this volume, chapter 3.3, §4):

Table 21: The Ezafe in Southeastern Kurmanjî (SEK)

	masc	fem	pl. (masc./fem.)
Definite	-(y)ē	-(y)a	-(y)ēd/-(y)ēt
Indefinite	-ū/-ē	-e/-a	
Demonstrative ezafe	yē	ya	yēt

One of the features distinguishing SEK from Standard K. and the rest of Kurmanjî is the use of the ezafe as a predicative element, rather than as part of the noun phrase. This phenomenon is discussed in MacKenzie (1961a: 205–208) and in Haig (2011); here we will only briefly outline it for Şemzînan (and SEK). Essentially it involves an ezafe which agrees in number and gender with its antecedent, but does not link that antecedent to some modifier; instead it introduces a verb phrase. Examples of this kind of usage are given below. (42) illustrates a clausal expression of possession (realized via the copula in Standard K.).

- (42) *min du bičük-ēt he-y*
 1SG.OBL two child-EZ.PL existent-NOT.ANALYZED
 ‘I have two children’

Predicates expressing location require a clause-final copula in Standard K., but in SEK the ezafe suffices:

- (43) *pīrežin-ek-ē* *li bin dīwar-i*
 old.woman-INDF-EZ.M at under wall-OBL.M
 'An old woman is at the base of the wall.'

In the present tenses of clauses with a nominal or adjectival predicate, the copula is combined with the ezafe.

- (44) *kuř-ek yī zīrek=e*
 boy=DEF EZ.M clever=COP.PRS.SG
 'The boy is clever.'

With finite verbs in the present tense, the ezafe expresses a progressive aspect:

- (45) *ber-ē xū da-yē kičik-a di-bēz-īt*
 direction-EZ.M self give.PST-3SG.GOAL girl-EZ.F IND-say.PRS-3SG
 '(S/he) looked at her (and saw that) the girl is saying: ...'

With past participles as main predicates, it forms the present perfect tense:

- (46) *hirč-ē yē xū lē da-y*
 bear-OBL.F EZ.M self at.him give.PST-PTCP
 'The bear has attacked him.'

- (47) *pīrežin-ē ser-ē da-na-y-e se*
 old.woman-OBL.F head-EZ.M PRV-put.PST-PTCP-DRCT on
ber-ek-ē
 stone-INDF-OBL.M
 'The old woman has put her head on a stone.'

Although it is fairly uncontroversial that these particles are etymologically identical with the ezafe, they are in fact not fully identical in form with the adnominal ezafes discussed further above, and there is some inconsistency in the forms used. In general, there seems to be a tendency for overgeneralization of the form -(y)ē, regardless of the gender of the antecedent (cf. (46) above). In the Yezidī dialects of Tur ‘Abdīn discussed in Bailey (2005), there is a similar neutralization of these tense-ezafes, leading to a uniform -ī.

3.2.2. Ezafe in WK

The ezafe forms and their functions constitute another domain where the WK dialect diverges remarkably from Standard K. Related forms are below:

Table 22: The Ezafe in Western Kurmanjî (WK)

	masc	fem	pl. (masc./fem.)
Definite	- <i>ı</i> / - <i>ē</i>	- <i>ē</i> / - <i>ɔ</i>	- <i>ē</i>
Indefinite	- <i>ı</i>	- <i>ē</i>	- <i>e</i>
Dem. ezafe	<i>ı</i>	<i>ɔ</i>	<i>ē</i>

With definite nouns the basic ezafe forms are -*ı* and -*ē*. The alternative forms -*ı* and -*ɔ*, masculine and feminine respectively, which are parallel to Standard K. forms, occur rarely and the conditions of their occurrence are not yet clear. The plural ezafe, as in Mardin, is a reduced form -*ē*. Thus, theoretically, in some cases of definite nouns, gender and number distinctions of ezafe are neutralized, illustrated in following examples:

- | | | |
|-----------------------|----------------------------|--|
| <i>ziman-ı/-ē mi</i> | ‘my tongue’ | (masc. sg.) |
| <i>mɔl-ē/-ɔ min</i> | ‘my home’ | (fem. sg.) |
| <i>sēv-ē mi</i> | ‘my apples’ | (pl.) |
| <i>mɔl-ē bɔv-ı te</i> | ‘your father’s home/house’ | (<i>mɔl</i> ‘home’: fem. sg.; <i>bɔv</i> ‘father’: masc. sg.) |

In indefinite nouns, however, the alternative forms are not used at all. Thus, the ezafe forms in indefinite nouns are the same with Standard K. in singular masculine and plural but differ from Standard K. in feminine, illustrated below:

- | | | |
|--------------------------|---------------------|-------------|
| <i>keçik-ek-ē rindik</i> | ‘a lovely girl’ | (fem. sg.) |
| <i>mērik-ek-ı gir</i> | ‘a big man’ | (masc. sg.) |
| <i>pisik-n-e řeš</i> | ‘(some) black cats’ | (pl.) |

The demonstrative or pronominal ezafe forms are substantially different from Standard K. and other dialects. A three way distinction (singular feminine and masculine, and plural) is preserved albeit with different forms.

- | | |
|----------------------------|------------------------------|
| <i>æv pisikɔ ɔ min=e</i> | ‘This cat (fem) is mine.’ |
| <i>æv xɔynɔ ı min=e</i> | ‘This house (masc) is mine.’ |
| <i>æv xɔynɔnɔ ē min=in</i> | ‘These houses are mine.’ |

The most distinctive feature of Elbistan WK morphosyntax is the obligatory use of what appears to be an ezafe, which cliticizes to the subject constituent of certain types of clauses. The examples below show the construction in copular clauses:

<i>æz-ī/-ē gir=im</i>	'I (masc./fem.) am big.'
<i>t-ī/-ē gir=æ</i>	'You (sg. masc./fem.) are big.'
<i>æw-ī/-ē gir=æ</i>	'She/he (masc./fem.) is big.'
<i>æm-e gir=in</i>	'We are big.'
<i>hūn-e gir=in</i>	'You (pl.) are big.'
<i>æw-ē gir=in</i>	'They are big.'

With negated non-verbal predicates, the ezafe particle attaches to the negation marker (Kömür 2003: 19):

<i>az nî birçî ma</i>	'I (male) am not hungry'	(<i>ne+i=nî</i>)
<i>az nê birçî ma</i>	'I (female) am not hungry'	(<i>ne+e=nê</i>)

The ezafe forms used with the singular pronouns correspond to the indefinite singulars (see above), while the plural indefinite ezafe is used only with pronouns of the first and second person plural. For the third person plural, the definite plural ezafe is used. These particles introduce gender distinctions into the first and second person singular of non-verbal clauses.

3.3. Adpositions in SEK

The system of adpositions in SEK differs from that of standard K. in several respects. Some of these are illustrated in (48), from Haig and Öpengin 2018 (glosses simplified). One of the three basic prepositions of standard K., *ji* 'from', is only present in SEK in a few formulaic expressions. In SEK, the sense of 'from' is covered by *li*, which also expresses 'in, at'. Example (48b) illustrates the preposition *li* in the sense of 'from'. Benefactives in SEK are expressed through the preposition *bo*, rather than the standard K. circumposition *ji ... ra*. This is shown in (48a). In addition to Standard K. *bi* 'with, through', SEK has also (*li*)*gel* or (*di*)*gel* 'with', as well as a further circumpositional *di ... da* 'inside', which is often reduced to the postpositional element (cf. 48c).

- (48) a. *tu hinde šīrē kīwīya nešēy bo min bīni*
you some milk.EZ.M goat NEG.can.2SG for me
SUBJ.bring.2SG
'Can't you bring some goat milk for me?'
- b. *ez dē šīrē kīwīya li kē_dere īnim*
I FUT milk.EZ.M goat from where (SUBJ)take.1SG
'Where shall I bring the goat milk from?'
- c. *du šēx malekē da čēnabin*
2 sheikh(PL) hous.IND.OBL in PRV.NEG.be.PL
'Two sheikhs in one house can't be.'

The prepositions *li*, *bi* and *di* are never realized as they are cited here; they are reduced to the consonantal element when preceding a vowel, and they show metathesis to *il*, *ib*, *id* preceding a consonant. They are thus realized as enclitics on whatever element precedes the prepositional phrase. This could be considered part of a general tendency to tolerate more complex syllable codas in SEK as opposed to dialects to the north and west. However, it also needs to be noted that simple prepositions are often completely elided (shown in parentheses), as seen in this proverb (49):

- (49) *cūn* (*ji/li*) *mirū=ye*, *hatin* (*ji/li*) *xudē*
 going (from) man=COP.3SG coming (from) God
 ‘Going is (from) man, returning (from) God.’

SEK makes use of simple ‘*bo X*’ construction, as in (48a) rather than Standard K. circumposition ‘*ji X re*’ for expressing benefactive; ‘*(li)gel X*’ rather than Standard K. ‘*bi X re*’ for comitative. Finally, the common Standard K. postpositional particle *re/ra* exists in SEK only in a circumposition *di ... řa* ‘through’ or its contracted pronominal form *tē řa* ‘through it’.

As in Standard K., there are also complex prepositions composed of a simple preposition and a local noun (cf. §3.5). In such combinations, the basic preposition is generally dropped, yielding what appears to be a new set of simple prepositions: *(li) se* ‘on’ (Standard K. *li ser*), *(li) nik* ‘beside’, *(li) bin* ‘under’.

3.4. Verbal morphology

3.4.1. SEK verbal morphology

3.4.1.1. Stem formation

As noted in §2.6, a number of Standard K. verbs have present stems consisting of either a bare consonant, or arguably, a consonant plus the short central vowel. In SEK, these verbs have what we refer to as “heavy” present stems, consisting of the initial consonant plus a vowel [æ] (orthographically <e>), or in the case of *xwarin*, a [o]. This is a feature shared in most of the SEK speech zone as well as in Central Kurdish. Examples of first person present tense forms of such verbs are in (50):

- | | |
|---------------------|-------------------|
| (50) <i>bi-ke-m</i> | <i>bi-be-m</i> |
| SUBJ-do.PRS-1SG | SUBJ-take.PRS-1SG |
| ‘I’ll do’ | ‘I’ll take’ |
|
 | |
| <i>di-de-m</i> | <i>di-xo-m</i> |
| IND-give.PRS-1SG | IND-eat.PRS-1SG |
| ‘I give’ | ‘I eat’ |

The present and past stem of a number of intransitive verbs (mostly “unaccusative”) in SEK have an extension *-(i)yē* (for present) and *-(i)ya* (for past); see Table 23. Furthermore, in a number of verbs, also shown in Table 23, the present and past stems of the verbs in SEK/Şemzînan are different from Standard K. and other dialects.

Table 23: Comparison of verb stems in Şemzînan (SEK) and Standard K.

Past		Infinitive +Gloss	Present		Infinitive +Gloss
Standard K.	Şemz.		Standard K.	Şemz.	
hişt-	hēla-	hēlan ‘leave’	-č- / -her-	-č-	čūn ‘go’
anī-	īna-	īnan ‘bring’	-gih- / -gihīž-	-geh-	gehiştin ‘reach’
axivī-	axiwt-	axiftin ‘speak’	-ē-	-hē-	hatin ‘come’
ēşīya-	ēşa-	ēşan ‘hurt’	-kişīn-	-kēş-	kēşān ‘pull’
kişand-	kēşa-	kēşan ‘pull’	-riž-	-rižiyē-	rižiyan ‘spill’
řizand-	řēt-	řētin ‘spill’	-řižīn-	-řež-	řežin ‘pour’
avēt-	howēt-	howētin ‘throw’	-āvēž-	-howēž-	howētin ‘throw’
ajot-	ha(w)jot-	ha(w)jotin ‘drive’	-x-	-ēx-	ēxistin ‘drop’
xist-	ēxist-	ēxistin ‘drop’	-girī-	-girī- / -giriyē-	giriyān ‘weep’

3.4.1.2. Preverb incorporation

In Standard K., there is a set of opaque preverbal particles such as *hil*, *řā*, *da*, which combine with verb stems to create new verbs. In the infinitive, they are usually written together with the stem as a single item. However, inflectional prefixes such as negation, or indicative/imperfective, are inserted between the preverb and the stem, as in Standard K. *ra-di-keve* ‘goes to sleep’, from *raketin* ‘go to sleep’. In SEK, however, negation and imperfective prefixes will often precede these preverbal particles, indicating full lexicalization of preverb+stem and the creation of a new stem. The same phenomenon is also found in the southernmost dialects of Southern Kurmanji, for example around Midyat; see Haig and Öpengin (2018), ex. (75) and accompanying discussion.

Examples (51a) and (51b) show preverb incorporation in Şemzînan (SEK). In (51a), the present indicative form of the verb *hel-(h)atin* ‘preverb+come’ (=‘rise’) has the indicative prefix preceding the preverbal element, and in (51b), both negation and indicative prefixes precede the preverbal element.¹¹

- (51) a. ſoj spēdē zū di-helē-t¹¹
 sun morning quick IND-rise.PRS-3SG
 ‘The sun rises early in the morning.’

¹¹ The verb form can be analysed as a contraction of *di-hel-hē-t* (IND-PRV-come.PRS-3SG). Cf. the corresponding standard K. form *hil-t-ē* (PRV-IND-come.PRS.3SG).

- b. *ez heta hēwari žī ne-di-řa-westīya-m*
 1SG until evening also NEG-IPFV-PRV-Stand.PST-1SG
 ‘I would not stop (working) until evening.’

Furthermore, in some highly lexicalized and frequent light verb constructions, the negation “prefix” can even occur on the leftmost edge of the verbal complex, preceding the non-verbal elements of the construction, as shown in (52).¹²

- (52) *min ne=ber-ē xū da-yē*
 1SG.OBL NEG=direction-EZ.M self give.PST-3SG.GOAL
 ‘I did not look at (him/her/it).’

3.4.1.3. Additional aspect distinctions

As already discussed under §3.2.1, an analytic “present progressive” can be expressed in SEK by using the ezafe, as illustrated in (53).

- (53) *ber-ē xo da-yē kičik=a di-bež-īt*
 direction-EZ.M self give.PST-3SG.GOAL girl=EZ.F IND-say.PRS-3SG
 ‘(S/he) looked (and saw that) the girl is saying (something).’

Other dialects of Kurmanjî in Turkey lack this possibility. Similarly, an alternative present perfect can be created using the ezafe in combination with past participles, as shown in examples (46) and (47). Note that these two tenses, present progressive and present perfect tense, constructed using the ezafe, are mostly restricted to affirmative and declarative clauses, as they are in Behdinî (cf. Haig 2011, Haig, this volume, chapter 3.3, §4). However, in Şemzînan the present progressive does lend itself to negation, illustrated in (54).

- (54) *axir tu yē na-xo-y*
 finally 2SG EZ.M NEG-eat.PRS-2SG
 ‘But you are not eating!’

3.4.2. Verbal morphology in WK

3.4.2.1. Person marking

The WK person marking system differs from Standard K. in that the copula forms of 2SG and 3SG are merged in -(y)ə. Similarly, the 2SG and 3SG verbal agreement suffixes are merged in -i [i]; see Table 24. In this manner, similar to the Mardin

¹² In Mêrd. dialect, a similar incorporation of preverbal particles can also be observed, cf. §4.2 in Haig and Öpengin (2018).

dialect, but unlike Şemzînan dialect of SEK, the person marking distinctions on verbs is reduced to three levels: 1SG – 2SG/3SG – 1PL/2PL/3PL.

3.4.2.2. *Gender marking in the predicate*

In WK, an ezafe particle attaches to the subject of non-verbal predicates, as shown in §3.2.2. A very similar usage also obtains in the present indicative, as in the following:

- (55) a. *aew-ī t-er-i*
 3SG-EZ.M IND-go.PRS-3SG
 ‘He goes’

- b. *aez-ē dæ-gē-m*
 1SG-EZ.F IND-reach.PRS-1SG
 ‘I (female) am arriving.’

- (56) a. *aez-ē te dæ-pē-m*
 1SG-EZ.F 2SG.OBL IND-wait.PRS-1SG
 ‘I (female) am waiting (for) you.’

- b. *t-ī dar-an xiš dæ-k-æ*
 2SG-EZ.M wood-PL.OBL PRV.cut IND-do.PRS-2SG
 ‘Are you (male) cutting the wood?’

However, it is yet to be confirmed whether clauses with full verbs in the past tenses allow for the subject to be further marked by the ezafe forms. It is absent on a number of past tense sentences in Çapar (2009). We conclude provisionally that ezafes attach to the subjects of present tense verbs, and to copular constructions irrespective of the tense, but we await a full account of the conditions on the use of the ezafe in other verbal constructions.

- (57) *aez-ē dæ-zɔn-im k=aew-ī læ vir bū*
 1SG-EZ.F IND-know.PRS-1SG that=3SG-EZ.M in here be.PST.3SG
 ‘I know that he was here.’
 (Çapar 2009: 63)

- (58) *[pisīk-n-e ūreš] geyrɔ-n*
 cat-INDF.PL-EZ black roam.around.PST-3PL
 ‘The black cats roamed around.’

- (59) *ře-yɔ xa šaš-miš kir*
 road-EZ.F self wrong-mIš do.PST.3SG
 ‘I lost my way’

The ezafe forms marking the subject in the present tense and copular constructions apply also to non-pronominal subjects. The resulting forms are (superficially) identical with oblique marked agents in past tense constructions of Standard K. and other dialects: *Musayîlæ viræ* ‘Musa-EZ.M (is) here’. With plurals, the ezafe applies regardless of whether the subject carries the plural oblique suffix, as in (60b), or does not carry it, as in (60a).

- (60) a. *pisîk-ē* *šîr* *væ-dæ-xɔ-n*
 cat-EZ.PL milk PRV-IND-eat.PRS-3PL
 ‘The cats are drinking milk.’
- b. *pisîk-ɔn-ē* *šîr* *væ-dæ-xɔ-n*
 cat-OBL.PL-EZ.PL milk PRV-IND-eat.PRS-3PL
 ‘The cats are drinking milk.’

Note that in this dialect, the plural oblique case suffix has been generalized to apply to nouns which in Standard K. would be in the direct case, as in (60b). However, it does not seem to have been fully reanalyzed as a generic plural suffix, since it does not systematically mark all the plural entities, hence the variation between (60a) and (60b).

With complex subject noun phrases, the particle occurs at the end of the subject phrase, as in (61):

- (61) a. *pisîk-n-e* *řeš-ē* *šîr* *væ-dæ-xɔ-n*
 cat-INDEF.PL-EZ black-EZ.PL milk PRV-IND-eat.PRS-3PL
 ‘The black cats are drinking milk.’
- b. *vî* *īlag-ɔ-y* *qilér=e*
 this shirt-PROX-EZ.M dirty=COP.3SG
 ‘This shirt is dirty.’

The same system apparently also applies to the future tense, according to the description in Kömür (2003: 18–20). The relevant forms are provided in Table 24. I assume that in the plural, the relevant clitic is uniform $=\hat{e}$, though this is not shown in the source. Note also the identical person marking suffixes in the second and third persons, as mentioned above. Unfortunately, we lack a detailed analysis of the verbal system of these dialects.

Table 24: Gender marking in Western Kurmanjî verbs ('send', present stem *şîn-*)

	PRESENT		FUTURE
1SG.M	<i>az=î</i>	<i>da-şîn-im</i>	<i>az=î ku bi-şîn-im</i>
1SG.F	<i>az=ê</i>	<i>da-şîn-im</i>	<i>az=ê ku bi-şîn-im</i>
2SG.M	<i>tu=yî</i>	<i>da-şîn-i</i>	<i>tu=yî ku bi-şîn-a</i>
2SG.F	<i>tu=yê</i>	<i>da-şîn-i</i>	<i>tu=yê ku bi-şîn-a</i>
3SG.M	<i>aw=î</i>	<i>da-şîn-i</i>	<i>aw=î ku bi-şîn-a</i>
3SG.F	<i>aw=ê</i>	<i>da-şîn-i</i>	<i>aw=ê ku bi-şîn-a</i>

3.4.2.3. Verbal negation

The negation prefix in past imperfective verb forms is *nɔ-*, identical with the negation prefix used in present indicative verbs, as in (62). In this feature, WK differs from Standard K., which uses the same negation prefix for all past tense verbs, and a different one for the indicative present. Furthermore, there is a distinct negation prefix for imperatives, *mæ-*, as in *mæ-wē* ‘do not say (it)’.

- (62) *gɔv-ɔ k=æz-ē læ mereš-ē wū-m min*
 time-EZ that=1SG-EZ in place.name-OBL be.PST-1SG 1SG.OBL
pir sēv nɔ-dæ-xɔr-in
 many apple NEG.IPFV-IPFV-eat.PST-3PL
 ‘When I was in Maraş, I would not eat so many apples.’

3.4.2.4. Turkish *miş*-verb forms in WK

A ubiquitous feature of all the western dialects is the massive influx of Turkish verb forms based on the Turkish perfect/evidential suffix -*miş*, combined with Kurdish light verbs, for example *an(l)amış kirin* ‘understand’ (Tk. *anlamış*), *qapatmış kirin* ‘close’ (Tk. *kapatmış*). The widespread use of such forms constitutes an important feature of these dialects as opposed to those of the southeast such as SEK, or SK, where at least in the speech of older speakers, such forms are rarely used (e. g. the extensive text material of Ritter, from Midyat region, or that of Nikitine from Şemzinan (in MacKenzie 1995) contain hardly a single form). But from WK, they are well attested in older sources (e. g. in the Kurmanjî texts of Le Coq 1903), and many are firmly established and phonologically adapted, as in (63).

- (63) a. *min ū-ya xa šaš-miš kir*
 1SG.OBL road-EZ.F self wrong-miš do.PST
 ‘I lost my way.’
 (Çapar 2009: 63)
- b. *insɔn-? dayan-miš na-b-ī ki*
 human-EZ stand-miš NEG-be.PRS-3SG PTCL
 ‘One cannot endure it.’

3.5. Issues in Western Kurmanjî (WK) syntax

Reflexive pronoun: In WK the reflexive pronoun in possessor function is generalized to be used in contexts where it is not controlled by a co-referential subject. It is thus used in much the same way as a 3SG oblique pronoun, as in (64).

- (64) a. *bɔv-ē xe čū-ye alwistan-ē*
 father-EZ.M REFL go.PST-DRCT¹³ place.name-OBL.F
 ‘His/her father has gone to Elbistan.’ (Standard K.: *bavê wî* ...)
- b. *ferq-a xe či=ye*
 difference-EZ.F REFL what=COP.3SG
 ‘What is its difference?’

A particle *ki*, homophonous to the particle also used in functions such as relative particle and subordinating conjunction, expresses the modality of “having the intention of doing something” (glossed as MOD), illustrated in (65).

- (65) *Sudi ew ki hatा türk baqɔl-ē har-in,*
 tomorrow 3PL MOD until turkish grocery-OBL.F go.PRS-3PL
ez=j̄t̄ ki v̄ē=rɔ har-im
 1SG=also MOD 3SG.OBL.F=POSTP go.PRS-3SG
 ‘Tomorrow they will go to the Turkish grocery store, I will also go with her.’
 (Çapar 2009: 78)

The *ki* particle can be used with the subject-marking *ezafe*, but it cannot be used with a future tense particle *-ē*. Note finally that the particle might originate from the auxiliary use of the verb *kirin* ‘do’ (present stem: *ki-*). In Standard K. and in central areas of Kurmanjî speech zone, as in SK, the conjugated form of the verb *kirin* is employed as the auxiliary in expressing the prospective aspect or the “immediate future”.

The conditionals in WK usually incorporate the Turkish clausal enclitic conditional marker *=se* to mark the verb of the protasis,¹⁴ as in (66). But the conditional conjunction *eger* and more widely the *ki* particle can also start the sentence.

- (66) *tu hat=se telafon-a mi ke*
 2SG come.PRS=COND phone-EZ.F 1SG.OBL do.IMPER.2SG
 ‘Call me if you come.’
 (Çapar 2009: 64)

Note that the *ki* relative/subordinating particle (Standard K. *ku*) is formally the same with the corresponding Zazaki (Haig 2001: 202; Paul 1998) and in all its functions it is usually a proclitic and reduced to the sole consonantal element.

The Standard K. adhortative particle *bila* does not exist in WK, a form *ma* is used in this function, as in (67).

¹³ This may be a present perfect formative, widely used in this dialect, rather than the directional particle. It is impossible to decide in this context (they cannot both be overtly realized on the same verb).

¹⁴ This is observed also for the geographically close Tunceli (Kr. Dersim) Kurmanjî in Haig (2006).

- (67) *tēlefon-a Domi ki-m, ma wer-i*
 phone-EZ.F proper.name do.PRS-1SG HORT come.PRS.SUBJ-3SG
 ‘I shall call Domi so that he comes’
 (Çapar 2011: 78)

4. Northern Kurdish in eastern Anatolia: summary of main contact issues

Northern Kurdish is spoken across most of eastern Anatolia, and has thus been exposed to contact influence from several different languages: In the southeast, it has co-existed for centuries with Neo-Aramaic and local varieties of Arabic, while probably the most important historical contact language in central and northeastern Anatolia would have been Armenian. More recently, Turkish has exerted considerable influence on all varieties of Kurmanjî, through Turkish-language mass media, compulsory schooling, military service, and large-scale migration to the main administrative centres, where representatives of the Turkish state tend to be concentrated.

Areally, Kurmanjî is split across the Mesopotamian zone and the Caspian/Caucasian zone, and variation in morphosyntax corresponds broadly to this north/south divide (Haig 2017). For example, the SEK dialects are firmly within the Mesopotamian zone, and here we find widespread use of non-canonical subjects with experiencer predicates, modal ‘want’, and expressions of possession (Haig 2006, 2017), a greater reliance on prepositions (§3.3), and a larger range of arguments that can occur post-predicatively (Haig 2015). The dialects of the north and west lack these features. This ties in with the general picture of Semitic influence in the southeast with a gradual fade-out northwards and westwards. But not everything fits this picture. It is quite unclear, for example, how areal considerations would be relevant in understanding the Western Kurmanjî constructions with the *ezafe* particle in the verbal domain (§3.2, 3.4).

In general, the core areas of Kurmanjî morphology show relatively little evidence of heavy structural borrowing (Haig 2007: 180). Most plausible candidates for contact-induced developments stem from phonology, lexicon, and syntax. The following list of candidate features for contact influence is not exhaustive, but merely illustrates some of those discussed in the literature:

1. Additional series of voiceless obstruents, presumably in part through Armenian influence (§2.2);
2. Pharyngealization, presumably through Semitic influence, but building on inherited features of the phonological system (Barry 2017);
3. Close similarities across the vowel systems in the languages of Anatolia (Haig 2017: 402);
4. Borrowing of Turkish conditional clitic =*ise* (cf. (66) from WK);

5. Use of Turkish *-mIš*-verb forms, incorporated into Kurdish complex predicates (cf. (59), (63) from WK);
6. Western Kurmanji dialects: numerals 11–19 follow Armenian pattern, reinforced by Turkish ('11' *dah-u-yek* etc. instead of *yānzdah* etc.), WK only;
7. Strategies for clause linkage (Matras 2002);
8. Common Anatolian clause-final copula construction (Haig 2017);
9. Borrowing of Turkish comparative particle *daha*;
10. Loss of the rule for reflexive binding with *xwe* 'self' (WK only, see Haig 2006, §3.5)

Previous research (Dorleijn 1996, Haig 2006, Haig 2007) has tended to focus on Turkish influence on Kurmanjî. While contemporary spoken Kurmanjî is undoubtedly heavily influenced by Turkish, it is important to consider the issue from a longer-term perspective. If we consider the situation of Kurdish prior to the founding of the Turkish state in 1923, there is little evidence of Turkish influence on much of Kurmanjî. Original texts recorded as late as the 1960's by Ritter (1971, 1976, see §5), show few traces of Turkish influence, either in lexicon or morphology. The same holds for most of what I have above termed Southeastern Kurmanjî (SEK), particularly in the far southeast of the country. There are still monolingual speakers of Kurdish in this region today, and we can reasonably assume that this was much more widespread a century ago. The texts compiled by Nikitine from this region in the early twentieth century reflect reasonably reliably the Kurdish at the time, and illustrate the general paucity of Turkish influence (see MacKenzie (1995) for a critical edition of one Nikitine's texts). In what we have termed Western Kurmanjî (WK), Turkish influence appears more deeply entrenched, and is evident in the texts of Le Coq (1903). In these texts, provided by speakers from Zincirli, west of today's Gaziantep, we already find the reflexive pronoun used as a general possessive marker, without being subject to the binding conditions that apply to Standard Kurmanjî, we find the numerals 11–19 in the Turkish/Armenian form rather than the inherited Iranian form, and a scattering of Turkish *mış*-verbforms. But none of these sources show anything approaching the massive Turkish influence (e. g. in terms of loan words, code-switching, Turkish-influenced syntax) that characterizes the casual speech of many Kurds today. While the data is still very sketchy, it seems reasonable to assume that up until the beginning of the twentieth century, Kurmanjî speech communities were able to foster and preserve their language over a vast region, and ensure unbroken transmission across generations. Levels of loan words in the basic vocabulary (see e. g. Haig and Öpentin 2014) are also low; despite centuries of co-existence, there are remarkably few clear cases of Armenian or Neo-Aramaic borrowings in the basic Kurmanjî lexicon, suggesting that Kurmanjî was not under any particular pressure from these languages (Kurdish loanwords in Neo-Aramaic on the other hand, are extremely numerous, see Khan 2007).

Considering today's situation, the most crucial difference to pre-republican times is the incomplete childhood acquisition of the full spectrum of grammatical and lexical oppositions of Kurdish. The destruction of traditional village networks, the intrusion of Turkish into the domestic sphere via the media, and most importantly, pre-school and primary school monolingual language policies, means early acquisition of Kurdish is interrupted, with concomitant loss or simplification of lexical and grammatical structure. Thus much of what is often considered "Turkish influence" (see e. g. Dorleijn 1996) can also be interpreted as the result of imperfect acquisition.

5. Short glossed text

The following text is an excerpt from the story 'The poor man, the snake, the Jew, and good fortune', recorded in the 1960's in one of the villages southeast of Midyat which, at that time, were still inhabited by Yezîdîs. The speaker was the Pîr of the local Yezîdîs, and the recording was made by a local Kurd, who was collaborating with the German Semitist Hellmut Ritter. They subsequently transcribed and published this and several other texts, together with a German translation, in Ritter (1976). Unfortunately, Ritter himself passed away in 1971, and the original magnetic tapes have never been recovered.

These stories represent one of the very few reliable records of spoken Kurmanjî from this period, and are typical of the oral tradition preserved in e. g. MacKenzie (1962), or Blau (1975), and discussed in Turgut (2012). Ritter's original transcription is phonetic rather than phonemic, and uses quite idiosyncratic symbols, making it rather inaccessible. In the version provided below, I have adapted it to the standard Kurmanjî orthography as outlined in Section 2 above, but the syntax remains as in the original, and dialectal features are noted where necessary.

- (68) *roj-ek¹⁵-ē kerk-ē¹⁶ wi li mil-ē wî ye COP.3SG*
 day-INDF-OBL.F yoke-EZ.M 3SG.OBL.M at shoulder-EZ.M 3SG.OBL.M
 'One day, his yoke is at his shoulder'

¹⁵ In the original transcription, the indefinite suffix is rendered with *-(i)k*, but I have standardized it throughout. Reduction of the indefiniteness suffix is a typical dialect feature of the Mardin region, where the suffix is realized as *[-(i)k]*, unlike Standard Kurmanjî *[-æk]*.

¹⁶ This word, transcribed in the original as *kärkē* (with ezafe), is translated into German as "*Holzgabel*", evidently a wooden artefact to enable a person to carry a load of wood on his or her shoulder. I have not been able to trace it in any of the Kurdish lexical sources known to me.

û bivr-ê wî li ser dest-ê wî ye
 and axe-EZ.M 3SG.OBL.M at on hand-EZ.M 3SG.OBL.M COP.3SG
 and his axe is in his hand.

- (69) *ji xwe=ra di-ç-e dar-a çîy-ê.*
 for self=POSTP IND-go-PRS.3SG wood-EZ.F mountain-OBL.M
 He is just going to the woods on the mountain.
- (70) *çû, dîn da-ye vaye mar-ek*
 go.PST.3SG sight give.PST.3SG=DRCT EXCL snake-INDF
 He went, and saw – what's that, a snake
ser-ê xwe der-êxist ba(ng) kiri=yê,
 head-EZ.M self out-put.PST.3SG calling do.PRF.3SG=3SG.OBL
 has popped out its head and called to him
- (71) *go kur-o! go ha!*
 say.PST.3SG fellow-VOC.M say.PST.3SG yes!
 saying: “fellow!”, he said “yes?”
- (72) *go ka were ez bêj-im=e te*
 say.PST.3SG PRT come.IMPER.SG 1SG say.PRS.SUBJ-1SG=DRCT 2SG.OBL
 He said: “won’t you come, that I may tell you (something)?”
- (73) *were cem mi(n)!*
 come.IMPER.SG to 1SG.OBL
 “come to me!”
- (74) *belengaz goti=yê go ya haywan-ê*
 poor.man say.PRF.3SG=3SG.OBL say.PST.3SG EXCL animal-EZ.M
xwedê god.OBL.M
 The poor man said to him, saying: “Oh creature of God,
- (75) *tu mar i û ez insan im*
 2SG snake COP.2SG and 1SG person COP.1SG
 you are a snake, and I am a human.
- (76) *ez=ê çawa b-êm=e cem te? [...]*
 1SG=FUT how SUBJ-come.PRS-1SG=DRCT to 2SG.OBL
 How should I come to you?”
- (77) *belengaz çû cem mér sekinî*
 poor.man go.PST.3SG to snake.OBL.M stop.PST.3SG
 The poor man goes up to the snake and waits.

- (78) *ya haywan-ê xwedê derd-ê te çi ye,*
EXCL animal-EZ.M god.OBL.M trouble-EZ.M 2SG.OBL what
COP.3SG
“Oh creature of God, what is your plight?
- (79) *tu çi ji mi(n) di-xwaz-î? gôti=yê go*
2SG what from 1SG.OBL IND-want.PRS-3SG say.PRF.3SG=3SG.OBL
say.PST.3SG
What do you want from me?” (He) said to him:
- (80) *ka bost-ik-ê ji dûv-ê mi(n) jê-bi-k-e!*
COMPL span-INDF-OBL.F from tail-EZ.M 1SG.OBL
from.it-SUBJ-do.IMPER-2SG
“cut off one span (unit of measure) from my tail!
- (81) *bost-a xwe bi-gir-e bi dûv-ê mi(n) jê-ke!*
span-EZ.F REFL SUBJ-take.IMPER-2SG through tail-EZ.M 1SG.OBL
Measure a span across my tail
and from.it-do.IMPER-2SG
and cut it off!”
- (82) *bê¹⁷ belê tu kêm santîn-k-î jê-k-e*
but 2SG less centimetre-INDF-OBL.M
from.it-SUBJ-do.IMPER-2SG
but if you cut off one centimetre too little
- (83) *ez=ê mal-a te xirab bi-k-im!*
1SG=FUT house-EZ.F 2SG.OBL ruined SUBJ-do.PRS-1SG
I will destroy your house!
- (84) *bi-hêl-im pirç-a ser-ê te bi-waş-e [...]*
SUBJ-let.PRS-1SG hair-EZ.F head-EZ.M 2SG.OBL
SUBJ-fall_out.PRS-3SG
I will cause the hair of your head to fall out.”

¹⁷ This reflects the original, though in Standard K. one might have expected *lê belê* here.

- (85) *ê go ya haywan-ê xwedê, ez na-wér-im*
 ez.M say.PST.3SG EXCL animal-EZ.M god.OBL.M 1SG
NEG-dare.PRS-1SG

The other one said: “Oh creature of God, I don’t dare
dûv-ê te jê-k-im
 tail-EZ.M 2SG.OBL from.it-SUBJ-do.PRS-1SG
 to cut off your tail.”

- (86) *go me-tirs-e baxt-ê xwedê ji te=ra*
 say.PST.3SG NEG-fear.IMPER-2SG fortune-EZ.M god.OBL.M from
you=POSTP
- He (the snake) said: “don’t be afraid, the fortune of God is with you.””

Abbreviations

1	first person	MOD	modality
2	second person	NEG	negation
3	third person	OBL	oblique
ADD	additive	PERF	perfect
ADP	adposition	PL	plural
AFF	affirmative	POSTP	postposition
COMPL	complementizer	PPRF	pluperfect
COND	conditional	PRES	present
COP	copula	PRF	perfect
DEM	demonstrative	PROX	proximal
DRCT	directional	PRS	present
EXCL	exclamative	PRT	particle
EZ	ezafe marker	PRV	preverbal particle
F	feminine	PST	past
FUT	future	PTCP	participle
HORT	adhortative	REFL	reflexive
IMPER	imperative	REL	relative
IND	indicative	SG	singular
INDEF	indefinite	SUBJ	subjunctive
IPFV	imperfective	VOC	vocative
M	masculine		

References

- Asatrian, Garnik. 2009. Prolegomena to the study of the Kurds. *Iran and the Caucasus* 13. 1–58.
- Bailey, Denise. 2005. A comparative study of grammatical relations in Northern Kurdish. Göttingen: University of Göttingen, unpublished MA thesis.
- Barry, Daniel. 2017. The emergence of pharyngeal sounds in Kurmanji Kurdish. Paper held at the 7th International Conference on Iranian Linguistics, Institute of Asian and African Studies, Lomonosov Moscow State University, 28–30 August, 2017.
- Bedir Khan, Emir & Roger Lescot. 1970. *Grammaire Kurde*. Paris: Adrien Maisonneuve.
- Blau, Joyce. 1975. *Le Kurde de 'Amādiya et de Djabal Sindjār: Analyse Linguistique, Textes Folkloriques, Glossaires*. Paris: Klincksieck.
- Bynon, Theodora. 1979. The ergative construction in Kurdish. *Bulletin of the School of Oriental and African Studies* 42. 211–224.
- Çapar, Fidan. 2009. Code-Switching und Sprachkontakt bei kurdischen Migranten aus Maraş. Frankfurt am Main: Johann Wolfgang Goethe-Universität, unpublished MA thesis.
- Dahl, Östen 2004. *The growth and the maintenance of linguistic complexity*. Amsterdam & Philadelphia: Benjamins.
- Dirêj, Evdila. 2011. *Kurdî Kurmancî*. Berlin: Dilop.
- Dorleijn, Margreet. 1996. *The decay of ergativity in Kurmanci*. Tilburg: Tilburg University Press.
- Haig, Geoffrey. 1998. On the interaction of morphological and syntactic ergativity: Lessons from Kurdish. *Lingua* 105. 149–173.
- Haig, Geoffrey. 2001. Linguistic diffusion in present-day East Anatolia: From top to bottom. In Alexandra Aikhenvald & R.M.W. Dixon (eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, 195–224. Oxford: Oxford University Press.
- Haig, Geoffrey. 2002. Noun-plus-verb complex predicates in Kurmanjî Kurdish: Argument sharing, argument incorporation, or what? *Sprachtypologie und Universalienforschung / Language typology and universals* 55(1). 15–48.
- Haig, Geoffrey. 2004. The invisibilisation of Kurdish: The other side of language planning in Turkey. In Stephan Conermann & Geoffrey Haig (eds.), *Die Kurden: Studien zu ihrer Sprache, Geschichte und Kultur*, 121–150. Schenefeld: EB.
- Haig, Geoffrey. 2006. Turkish Influence on Kurmanjî: Evidence from the Tunceli Dialect. In Lars Johanson & Christiane Bulut (eds.), *Turkic-Iranian Contact Areas: Historical and Linguistic Aspects*, 283–299. Wiesbaden: Harrassowitz.
- Haig, Geoffrey. 2007. Grammatical borrowing in Kurdish (Northern Group). In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 165–183. Berlin & Boston: De Gruyter Mouton.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A Construction Grammar approach*. Berlin & Boston: De Gruyter Mouton.
- Haig, Geoffrey. 2011. Linker, Relativizer, Nominalizer, Tense-Particle: On the Ezafe in West Iranian. In Yap Foong, Karen Grunow-Härsta & Janick Wrona (eds.), *Nominalization in Asian Languages: Diachronic and Typological Perspectives, vol. 1, Sino-Tibetan and Iranian Languages*, 363–392. Amsterdam & Philadelphia: Benjamins.
- Haig, Geoffrey. 2015. Verb-Goal (VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 407–425. New York: Gorgias Press.

- Haig, Geoffrey & Necle Bulut. 2017. Hearing, speaking, looking and learning: A pilot study of lexical variation across Kurmanjî Kurdish. *Wiener Jahrbuch für Kurdische Studien* 5. Special edition: Katharina Brizić, Agnes Grond, Christoph Osztovics, Thomas Schmidinger & Maria Six-Hohenbalken (eds.), Sprache-Migration-Zusammenhalt: Kurdisch und seine Diaspora. 99–109.
- Haig, Geoffrey & Ergin Öpentin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 3. 99–122.
- Haig, Geoffrey & Ergin Öpentin. 2015. Gender in Kurdish: Structural and socio-cultural aspects. In Marlis Hellinger & Heiko Motschenbacher (eds.), *Gender across Languages*, vol. 4, 247–276. Amsterdam & Philadelphia: Benjamins.
- Haig, Geoffrey & Ergin Öpentin. 2018. Kurmanjî Kurdish in Turkey: Structure, varieties, and status. In Christiane Bulut (ed.), *Linguistic Minorities in Turkey and Turkic speaking minorities of the peripheries*, 157–230. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1977. Zur Phonologie des Kurdischen in der Türkei. *Studien Zur Indologie und Iranistik* 3. 84–106.
- Khan, Geoffrey. 2007. Grammatical borrowing in North-eastern Neo-Aramaic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*. 197–214. Berlin & Boston: De Gruyter Mouton.
- Khan, Geoffrey. 2008. *The Neo-Aramaic dialect of Barwar*, vol. 1, Grammar. Leiden: Brill.
- Kömür, Mehmet. 2003. *Taw dilo: Akçadağ-Elbistan ağızıyla Kürt folkloründen seçimeler* [Taw dilo: Selections from Kurdish folklore from the Akçadağ-Elbistan dialects]. İstanbul: Weşanên Enstîtuya Kurdî ya Stenbolê.
- Lewendî, Mehmûd. 1997. Kurdên Şêxbizînî [The Şêxbizînî Kurds]. *Bîrnebûn* 3. 78–98.
- Le Coq, Albert von. 1903 [2003]. *Kurdische Texte: Kurmângî-Erzählungen und -Lieder, nebst einer Zazâ-Erzählung. Teil I und II*. Berlin. [Reprint Amsterdam: APA-Philo Press].
- MacKenzie, David. 1961a. *Kurdish dialect studies I*. Oxford: Oxford University Press.
- MacKenzie, David. 1961b. The origins of Kurdish. *Transactions of the Philological Society* 60(1). 68–86.
- MacKenzie, David. 1962. *Kurdish dialect studies II*. Oxford: Oxford University Press.
- MacKenzie, David. 1995. The Kurdish of Mullâ Sa'îd Shamdinânî. *The Journal of Kurdish Studies* 1. 1–27.
- Matras, Yaron. 1997. Clause combining, ergativity, and coreferent deletion in Kurmanjî. *Studies in Language* 21. 613–653.
- Matras, Yaron. 2002. Kurmanji complementation: Semantic-typological aspects in an areal perspective. *Sprachtypologie und Universalienforschung / Language typology and universals* 55. 49–63.
- Matras, Yaron et al. 2016. The Dialects of Kurdish. Web resource, University of Manchester. <http://kurdish.humanities.manchester.ac.uk/> (accessed 10 November 2017).
- McWhorter, John H. 2005. *Defining Creole*. Oxford: Oxford University Press.
- Oberling, Pierre. 2008. Khorasan i. Ethnic groups. *Encyclopædia Iranica*. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/khorasan-1-ethnic-groups> (accessed 16 September 2017).
- Öpentin, Ergin & Geoffrey Haig. 2014. Regional variation in Kurmanjî: A preliminary classification of dialects. *Kurdish Studies* 3. 143–176.
- Paul, Ludwig. 1998. The position of Zazaki among West Iranian languages. In Nicholas Sims-Williams (ed.), *Proceedings of the 3rd European Conference on Iranian Studies*, 163–176. Wiesbaden: Reichert.

- Paul, Ludwig. 2009. Zazaki. In Gernot Windfuhr (ed.), *The Iranian languages*, 545–586. London & New York: Routledge.
- Ritter, Helmut. 1971. Kurmanci-Texte aus dem Tur'abdin. Teil I. *Oriens* 21–22. 1–135.
- Ritter, Helmut. 1976. Kurmanci-Texte aus dem Tur'abdin Teil II. *Oriens* 25–26. 1–37.
- Schroeder, Christoph. 1999. Attribution in Kurmancî (Nordkurdisch). In Karl Wagner & Wolfgang Wildgen (eds.), *Studien zur Phonologie, Grammatik, Sprachphilosophie und Semiotik*, 43–63. Bremen: Institut für Allgemeine und Angewandte Sprachwissenschaft.
- Thackston, Wheeler M. 2006. *Kurmanjî Kurdish: A reference grammar with selected readings*. Unpublished Manuscript, Harvard University. <http://www.fas.harvard.edu/~iranian/Kurmanji/> (accessed 15 January 2015).
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Turgut, Lokman. 2012. *Mündliche Literatur der Kurden in den Regionen Botan und Hekarî*. Berlin: Logos.

2.4. The Arabic dialects of eastern Anatolia

Stephan Procházka

Vernacular Arabic is spoken in a number of non-connected parts of southeastern Anatolia from the city of Urfa in the west to the city of Siirt in the east.¹ The region thus roughly covers the part of Mesopotamia that lies within the borders of modern-day Turkey.

1. General remarks

1.1. Genetic affiliation

Vernacular Arabic belongs to the Semitic branch of the Afro-Asiatic language family. The varieties of Arabic spoken in eastern Anatolia can be divided into two large groups. The eastern, sedentary-type dialects are spoken mainly in the provinces of Mardin, Batman, Muş and Siirt, and belong to the Anatolian branch² of the Mesopotamian dialect group. They are often called *qəltu* dialects, from the verb *qəltu* 'I spoke'. The western dialects, which are of the Bedouin type, form the northernmost branch of the Arabian dialect group and constitute a sub-group of the so-called *Šāwī* Arabic, i. e. the language of the sheep- and goat-breeding semi-nomads of the Syrian Desert.

1.2. Brief history of the speech community

The history of settlement of eastern Anatolia has yet to be written. The Arab presence in the region goes back to the Parthian period when they organized border states at Hatra and Urfa (Edessa). Under the Sasanians members of the Rabī'a and Taglib tribes had settled around Nuṣaybīn (Nasibin) and Beth Arbaye (Morony 2005: 215, 221). Despite this strong presence of Arabs in pre-Islamic times, it can be assumed that most ancestors of today's Arabic speakers arrived in the aftermath of the Muslim conquest of the region in the middle of the 7th century CE. Political history indicates that the core of the Arabic-speaking sedentary population, i. e. the speakers of the *qəltu* dialects, may have settled in SE Anatolia during the 7th and early 10th centuries, when the region was under the stable control of Arab dynasties

¹ For details see the maps in Jastrow (1978), Procházka (2003: 88) and Procházka (2018) as well as the lists of villages in Andrews (2002).

² A detailed classification of Anatolian Arabic is given in Jastrow (1994: 121) and Jastrow (1978: 3–23).

like the ḥAbbāsids and Hamdānids.³ This is corroborated by the close linguistic relationship of the Eastern Anatolian *qəltu* Arabic with the dialects spoken in northern Iraq and, albeit less distinctive, in northern Syria.⁴ However, over the centuries many changes due to emigration, displacement, and resettlement took place. The *qəltu*-Arabic-speaking communities, with the exception of Mardin and its surroundings, have become isolated language islands amidst Kurdish- and — restricted to the towns — Turkish-speaking majority populations.⁵ This picture is clearly reflected in the “dialectal landscape”: the Mardin dialect group is much more conservative, and thus closer to Northern Iraqi Arabic, than the dialects of the “outer circle” of Arabic-speaking enclaves (Jastrow 2006–2007: 64). A good example of how volatile ethnic and religious identities can be is the group called the Mḥallamīye, who live between Mardin and Midyat.⁶ Their self-designation most likely goes back to the name of a Kurdish tribe and they may have been Christians until around 1600 (Benninghaus 2002: 179). Thus they originally spoke Aramaic and were Arabized only after they became Muslim (Jastrow 2004: 99). Another example of shifting identities in the region is the saying of a man from Mardin recorded in Grigore (2007: 34): *ana tərki-ana, abūy kərdi-we w əm̩mi Sarabīye-ye* ‘I am a Turk, my father is a Kurd, and my mother is an Arab.’

It has to be added that in all but the remote mountain regions regular, although often hostile, contacts with the Arab nomads of the Syrian Steppe have always existed.

In the Harran-Urfa area there are seven important Arabic-speaking tribes (with numerous sub-tribes) whose presence in the region north of the Euphrates probably goes back to the very first centuries of Islam (Oppenheim 1939: 222). The immigration of Arab tribes to the region reached a peak in the late 10th century (Heidemann 2002: 32), but migration to and from the Jazeera happened up to the First World War. Many of them were semi-nomads until the middle of the 20th century and therefore had close contacts with their fellow tribes-people in the

³ Cf. *Encyclopedia of Islam (EI)*, s. v. Anadolu, Harrān, Hıṣn Kayfā, Mardin, Siğird. It is worth mentioning that the eastern branch of the line of fortifications against the Byzantines, the *tugūr al-Ğazīra*, was “only so-called because the *ribāṭs* there are manned by the men of Mesopotamia” (*EI*, s. v. *al-Thughūr*).

⁴ Local tradition, however, does not always coincide with these assumptions. For example, the two Arab tribes in the village of Tillo trace themselves back to immigrants who settled in the region only in the 14th and 17th centuries, respectively (Lahdo 2010: 26).

⁵ The more or less complete isolation of all Anatolian Arabs from the rest of the Arabic-speaking world took place only after 1923, when it became almost impossible for them to cross the borders into Syria and Iraq. This situation only changed in the late 1990s.

⁶ Tens of thousands of them emigrated to Lebanon after WW I, where they are now called Mardilli and regarded as Kurds. Kern (2015) presents a very good overview on the different Mḥallami identities.

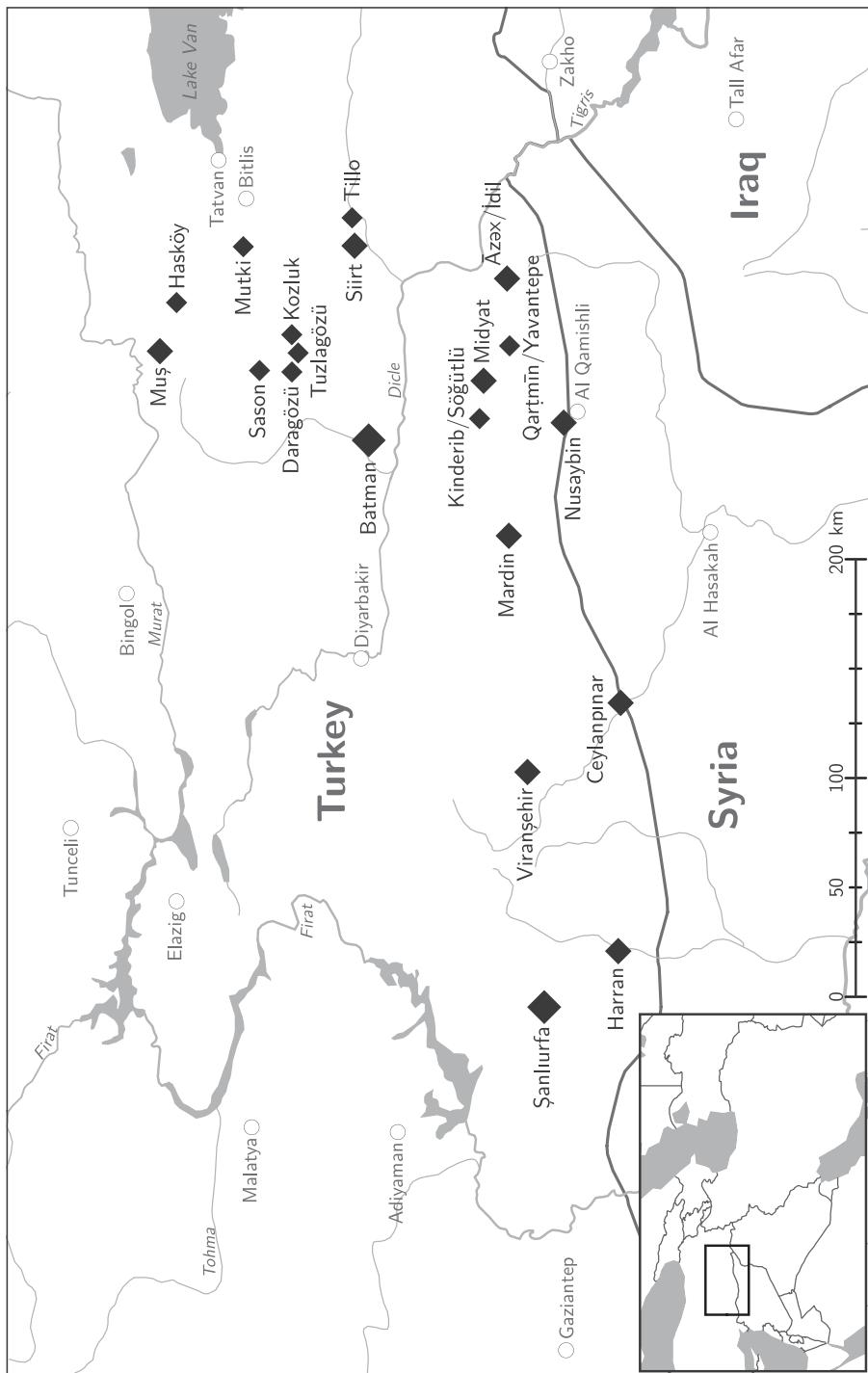


Figure 1: Location of Arabic dialects discussed in this chapter

Syrian Desert —which explains the relatively few linguistic differences between these groups.⁷

1.3. Current size and status of the speech community: religious affiliation

The last official census of the Republic of Turkey which gathered data relevant to languages dates back to 1965. Hence all numbers given in the following are based on more or less rough estimates. The *q̠altru*-Arabic speakers of eastern Anatolia number approximately 620,000 people⁸, and the speakers of Harran-Urfa Arabic can be estimated to be around 150,000⁹. The latter are all Sunni Muslims. Among the other group were significant Christian and Jewish minorities until a few decades ago. Today, there are no Jews in the region, and many of the Christians have also left, particularly those who were living in villages or in the smaller towns.¹⁰ Thus the overwhelming majority of the remaining Arabic-speaking population is Sunni Muslim.

1.4. Sociolinguistic situation

In most of the area covered by this article Kurdish is the dominant language of daily life, and Turkish is the official language (Talay 2006–2007: 180). Today, almost all of the Arabic speakers have a very high proficiency in Turkish, which dominates education, business life, and the media. Turkish is also the language exclusively used for written communication by Anatolian Arabs. Knowledge of Kurdish has decreased among the younger generation in favor of Turkish, and depends very much on the geographical setting. Many inhabitants of mixed Arabic-Kurdish villages speak Kurdish and Turkish and thus are trilingual.¹¹ In larger towns, particularly in Urfa, few Arabic speakers rather than craftsmen or merchants, are fluent in Kurdish. Contacts of Arabs with speakers of Aramaic (particularly *Türōyō*) may have been relatively intense, especially in the region west of Tur Abdin. Due to the rapid decline of Aramaic in Turkey¹², such contacts have ceased to play any but a very marginal role in the region's sociolinguistic setting.

⁷ One has to bear in mind that until the 1920s there were no political borders between what is today Syria, Iraq, and Turkey.

⁸ This number is given on the “Ethnologue” website www.ethnologue.com/language/ayp (accessed 8 July 2016).

⁹ Cf. Procházka (2003: 76).

¹⁰ Cf., e. g. Wittrich (2001: 4).

¹¹ Cf. Benninghaus (2002: 179) for the M̄hallami; and Lahdo (2009), which is a good case study on the situation in the village of Tillo in the Siirt Province.

¹² Only 15,000 speakers of Aramaic were left in Turkey in 2014 (data according to www.ethnologue.com/language/tru, accessed 14 July 2016).

On the whole, both branches of Eastern Anatolian Arabic are still vigorous, non-endangered varieties of Arabic with the following caveat: dialects spoken in small enclaves whose speakers have few contacts with other Arabic speakers must be considered highly endangered. Many Arabic speakers in such places suffer from both poor economic conditions and increasing Kurdish nationalism and therefore leave their homes. This is particularly true of Christians but has been reported for Muslim villages as well.¹³ But larger communities such as the Arabs of Mardin may also have difficulties in resisting the social and economic pressures that are resulting in a gradual shift to Turkish among the young people, many of whom recent studies have shown lack full proficiency in Arabic (Grigore 2007: 34–36). Such developments are also observed among the youngest generation of speakers of Harran-Urfa Arabic.

In general, the prestige of local Arabic is rather low. For many decades minority languages were completely ignored if not actively suppressed by the Turkish authorities. This attitude only changed around the year 2000 with an ever-increasing public and academic interest in local spoken languages. These developments may strengthen the self-confidence of the Arabic communities and foster efforts towards language preservation among their younger generation. There are increasing activities on internet sites as well as on local radio and TV stations to promote local varieties of Arabic. Another event which has changed the sociolinguistic setting is the massive immigration of Syrians who fled from the civil war that began in 2011. Thus, speakers of Urfa Arabic in particular have begun to adapt their speech to the dialect of the Syrian migrants and increasingly use typical words and phrases of Syrian Arabic.

As most Arabic speakers of the region in question are illiterate in Arabic, the local vernacular is usually not written. There are, however, numerous internet sites where one can find single phrases, songs, and proverbs in vernacular Arabic written in Latin script.¹⁴

2. Linguistic description

The huge differences between the two varieties of Arabic here under consideration, which are mutually not understandable,¹⁵ make it necessary to split many

¹³ According to Lahdo (2009: 112) the number of Arabic speakers in Tillo decreased from 3,000 to less than 1,000 during the five years from 2000 to 2005.

¹⁴ A good example is <http://www.mahalmi.com/default.asp?t=23> (accessed 28 February 2013) and the “Dictionary” of the dialect found on http://www.haldeh.com/index.php?option=com_content&view=category&id=44&Itemid=119 (accessed 8 July 2016).

¹⁵ Communication between local Arabs and the Arabic-speaking farm workers from the Mardin area who annually come to the Plain of Harran for harvest is usually in Turk-

chapters of the following linguistic sketch. For the *qəltu* dialects, this article will mainly focus on the dialect of the region of Mardin, where by far most of the *qəltu*-Arabic speakers live. However, outstanding linguistic features of other than this region will be referred to.¹⁶ For the sake of convenience the western, Bedouin, dialects will be called Harran-Urfa Arabic, the eastern, sedentary, dialects, *qəltu* Arabic when referring to the whole group, otherwise Mardin Arabic.

2.1. Phonology

2.1.1. Consonants

The following table includes the **consonantal phonemes** of Mardin and Harran-Urfa Arabic. Sounds marked by * occur in Mardin Arabic only.

Table 1: Consonant inventory

	Bilabial	Labio-dental	Inter-dental	Dental-alveolar	Palatal	Velar	Uvu-lar	Pharyngeal	Glottal
Plosive									
voiceless	p			t t [t]		k	q		(?)
voiced	b			d		g			
Affricate					č [tʃ̪]				
voiceless					č [tʃ̪]				
voiced					ğ [dʒ̪]				
Fricative									
voiceless	f	t [θ̪]	s s [s]	š [ʃ̪]	x		h [h]		h
voiced	*v	d[ð̪]	z	*ž [z̪]	ğ [ɣ̪]		h		
Nasal	m			n					
Lateral				l l [l̪]					
Trill				r r [r̪]					
Approximant	w				y				

ish. According to informants the situation is different in the easternmost parts of the province of Urfa. The Arabic vernacular of the regions of Viranşehir and Ceylanpınar reputedly shows more influence from the neighboring *qəltu* dialects.

¹⁶ Detailed descriptions of dialects from the whole region are found in Jastrow (1978), which is the classic monograph on *qəltu* Arabic. Other comprehensive studies are Grigore (2007) on Mardin, Jastrow (1973) on Daragözü, Lahdo (2010) on Tillo, Sasse (1971) on Mhallamîye, Talay (2001) and (2002) on Hasköy, Wittrich (2001) on Azex/İdil, Akkuş (2016) on the Mutki-Sason region.

The sound /g/ is only marginal in Harran-Urfa Arabic because it is usually reflected as /q/ (*qanam* ‘sheep’)¹⁷. There are several sub-groups of *qəltu* Arabic where the interdental fricatives (*d*, *t*, *θ*) have shifted to dental stops (*d*, *t*, *d*), sibilants (*z*, *s*, *z*), or labio-dental fricatives (*v*, *f*, *y*), respectively (Jastrow 1978: 34–39).

A very common phenomenon of *qəltu* Arabic is the devoicing of voiced consonants in final position: e. g. *p̪id* > *p̪it*# ‘hand’, *b̪id* > *b̪it*# ‘white (PL)’ (Jastrow 1978: 98); and (Siirt) *g̪ēmeʃ* > *g̪ēmeh*# ‘mosque’.¹⁸ The same phenomenon exists in Urfa Arabic, e. g. *aṭluʃ* > *aṭluh*# ‘I ascend’.

2.1.2. Vowels

Mardin Arabic possesses two short (/ə/ and /a/) and five long (/ā/, /ē/, /ī/, /ō/, /ū/) vowel phonemes as well as two diphthongs, /ay/ and /aw/ (Jastrow 1978: 81–84). The vowels [ē] and [ō] are often products of the lowering of /i/ and /u/ when preceding a velarized or back consonant: *malēh* < *malīh* ‘good’, *tyōr* < *tyūr* ‘birds’ (Grigore 2007: 77). Diachronically seen, the vowel /ē/ has also emerged from /ā/ by the shift of /ā/ > /ē/ conditioned by a short or long *i* in an adjacent syllable: e. g. Mardin *gmēl* < *gimāl* ‘camels’.¹⁹

Harran-Urfa Arabic has the same five long vowels²⁰ as Mardin Arabic but three short vowels, /a/, /i/ and /u/. The phonemic status of /i/ and /u/ is weak and only attested by a handful of minimal pairs.

2.1.3. Syllable structure²¹

In both groups the following types of syllables occur: Cv, CvC, Cv̄, Cv̄C, CCv, CCv̄, CCv̄C, and CCv̄C. The syllables CvCC, Cv̄CC,²² and CCv̄CC are found among the *qəltu* dialects only, since in Harran-Urfa Arabic -CC# is normally realized as -CvC#.²³ In the *qəltu* dialects there is a tendency to articulate initial CC- with an epenthetic, non-phonemic /(?ə)/: e. g. [əftáh ~ ʔəftáh ~ ftáh] ‘open!’

¹⁷ Whereas OA /q/ is reflected as /g/: e. g. *bugar* ‘cows’.

¹⁸ The devoicing of /ʃ/ does not occur in Mardin Arabic (Jastrow 1978: 43).

¹⁹ For a detailed description of this phenomenon called *?imāla* cf. Jastrow (1978: 65–69).

²⁰ The two long vowels /ē/ and /ō/ are products of the monophthongization of OA /ay/ and /aw/.

²¹ The syllable structure of the Anatolian *qəltu* dialects is presented in detail by Jastrow (1978: 85–95).

²² The syllable C₁v̄/v̄C₂C₃ does not occur with C₃ = *m*, *n*, *l*, *r*: e. g. **fa-tahl-ki* > *fataḥəlkı* ‘he opened (for) you’, but *xalf-na* ‘behind us’ (Jastrow 1978: 88).

²³ In the *qəltu* dialects only if the last C = *m*, *n*, *l*, *r*: e. g. *šahr* > *šahər* ‘month’ (see also preceding footnote).

(Jastrow 1978: 87).²⁴ This is also true of Harran-Urfa Arabic, but there forms like *trāb* ‘soil’ and *ghawa* ‘coffee’²⁵ can also be heard.

In all Anatolian *qəltu* dialects /ə/ (< /i/, /u/) is elided in unstressed open syllables unless this elision causes a cluster of CCC: e. g. Mardin *yqēfəd* ‘he sits down’, *yqēfdūn* ‘they sit down’; but *yəktəbūn* ‘they write’.²⁶ If, however, the vowel /ə/ goes etymologically back to /a/, it is not elided (Jastrow 1978: 62): e. g. *šərəbna* ‘we drank’ < Old Arabic (= OA) *šaribnā*.

In Harran-Urfa Arabic too, /i/ and /u/ are elided in unstressed open syllables: e. g. *yākul* ‘he eats’, *yāklūn* ‘they eat’. However, both vowels are retained if they go back to OA /a/: e. g. *zimān* < *zamān* ‘time’. Certain patterns are subject to sound changes if suffixes are attached. This is especially true for words of the structure *CaCaC*, which show regular re-syllabification regardless if the word in question is a verb or a noun. For instance, *ḍarab* ‘he hit’ vs. *ḍribat* ‘she hit’; *hağal* ‘partridges (coll.)’ vs. *ḥgile* ‘partridge’, *maras* ‘rope’ vs. *mrise* ‘thick rope’.²⁷

2.2. Nominal morphology

2.2.1. Inflection

As in all Arabic vernaculars there are no case markers in Eastern Anatolian Arabic. Nouns are inflected according to the three numbers singular, dual, and plural. The singular is not marked. The suffix of the dual is *-ayn* (Mardin) and *-ēn* (Harran-Urfa). The plural is formed by suffixes (in particular *-āt* and *-īn*) or by modification of the word pattern.

Examples from Harran-Urfa Arabic:²⁸

	time, instance	ram	door
SG	<i>nōba</i>	<i>čabiš</i>	<i>bāb</i>
DUAL	<i>nōbtēn</i>	<i>čabšēn</i>	<i>bābēn</i>
PL	<i>nōbāt</i>	<i>čbāš</i>	<i>būb</i>

²⁴ Only in the Kozluk-Sason group is the epenthetic vowel inserted: e. g. **ftah* > *fətáh* ‘open!’ (Jastrow 1978: 87).

²⁵ Unless otherwise indicated the examples from Harran-Urfa Arabic are from unpublished own research.

²⁶ Exceptions from that rule are found in the Kozluk-Sason group (Jastrow 1978: 88).

²⁷ Cf. Procházka (2003: 77).

²⁸ In Harran-Urfa Arabic, the plural suffix *-īn* is almost exclusively used with adjectives and participles.

2.2.2. Pronouns

The following table shows the independent and bound personal pronouns of Mardin and Harran-Urfa Arabic (from Procházka 2006–2007: 125 and Jastrow 1978: 127–131; Jastrow 2006: 90).

Table 2: Personal pronouns

		independent forms	bound forms	
	Mardin	Harran-Urfa	Mardin after C—after V	Harran-Urfa after C—after V
1 SG	ʔana	ʔāni	-i/-ni — -yi/-ni	-i/-ni — -ya/-ni
2 SG M	ʔənta	ʔinte	-ək — -k	-ak — -k
2 SG F	ʔənti	ʔinti	-ki — -ki	-ič — -č
3 SG M	hūwe	huwwa	-u — -hu	-u — ү ²⁹
3 SG F	hīye	hiyye	-a — -ha	-ha — -ha
<hr/>				
1 PL	nəħne	ʔiħna	-na — -na	-na — -na
2 PL M	ʔəntən	ʔintu	-kən — kən	-kum — -kum
2 PL F		ʔintin		-čin — -čin
3 PL M	hənne	humħa	-ən — -ħən	-hum — -hum
3 PL F		hinne		-ħin — -ħin

A salient feature of the Anatolian *qəltu* dialects (with some exceptions in the Kozluk-Sason group) is the sound *n* in several plural pronouns as well as in inflection morphemes of the perfect. As can be seen from the corresponding forms in Harran-Urfa Arabic, the *qəltu* forms (as in *ħənne*, *bayt-ən*, *ʔəntən*, *bayt-kən*) coincide with the respective feminine pronouns whereas Harran-Urfa masculine pronouns have *-m* instead.³⁰ It should be mentioned that Harran-Urfa Arabic exhibits gender distinction in all plural forms except the first person in both pronouns and verbs (see below).

²⁹ The pronominal suffix is indicated by the lengthening and stressing of the final vowel.

³⁰ For the pronouns that show an *-n*, Aramaic substratum influence has been suggested. The pros and cons of such a development are discussed in detail in Arnold and Behnstedt (1993: 75–79), Diem (1979: 44), Owens (2006: 244–245), Weninger (2011: 749).

2.3. Verbal morphology

2.3.1. Inflection

Finite verbs are conjugated for **perfect** and **imperfect**. In *qəltu* Arabic two different vowel patterns exist in the basic stem (Form I): *CaCaC* and *CəCəC* for the perfect and *-CCaC*, and *-CCəC* for the imperfect. Harran-Urfa Arabic possesses two stems for the perfect, *CCiC* and *CaCaC* (the latter with the phonological variants *CiCaC* and *CuCaC*), and three for the imperfect, *-CCaC*, *-CCiC*, and *-CCuC*.

Table 3: Paradigms for the perfect and imperfect in Mardin and Harran-Urfa

	Perfect		Imperfect	
	Mardin	Harran-Urfa	Mardin	Harran-Urfa
1 SG	šəṛəb- <i>tu</i>	šrib- <i>it</i>	ʔa-šrab	ʔa-šrab
2 SG M	šəṛəb- <i>t</i>	šrib- <i>it</i>	ta-šrab	ti-šrab
2 SG F	šəṛəb- <i>ti</i>	šrib- <i>ti</i>	ta-šrab- <i>īn</i>	ti-šrab- <i>īn</i>
3 SG M	šəṛəb	šrib	yə-šrab	yi-šrab
3 SG F	šərb- <i>ət</i>	širbat	ta-šrab	ti-šrab
<hr/>				
1 PL	šəṛəb- <i>na</i>	šrib- <i>na</i>	nə-šrab	ni-šrab
2 PL M	šəṛəb- <i>tən</i>	šrib- <i>taw</i>	ta-šrab- <i>ūn</i>	ti-šrab- <i>ūn</i>
2 PL F		šrib- <i>tan</i>		ti-šrab- <i>in</i>
3 PL M	šərb- <i>u</i>	širb- <i>aw</i>	yə-šrab- <i>ūn</i>	yi-šrab- <i>ūn</i>
3 PL F		širb- <i>an</i>		yi-šrab- <i>in</i>

In the *qəltu* dialects, the final *-n* of the inflection morphemes *-īn* and *-ūn* is dropped when a pronoun suffix is attached: e. g. *yišrabū-hu* ‘they drink it’ (Jastrow 2006: 93). In Harran-Urfa Arabic the *-n* is retained: e. g. *yšūfūn-ha* ‘they see her’.³¹

³¹ In Harran-Urfa Arabic, the final *-n* of the plural morpheme is elided when an indirect object pronoun is attached to a participle: e. g. *gāylīn* ‘they have told’, but *gāylī-li* ‘they have told me’.

Imperative

Table 4: Imperatives in Mardin and Harran-Urfa

	Mardin	Harran-Urfa
2 SG M	šrab	ʔišrab
2 SG F	šrab-i	ʔišrab-i
2 PL M	šrab-u	ʔišrab-aw
2 PL F		ʔišrab-in

Mardin and Harran-Urfa Arabic possess a special particle which serves to emphasize the imperative, *dē* ~ *də-* in Mardin and *di-* ~ *du-* in Harran-Urfa: e. g. Mardin: *də-qūmu* ‘stand up (PL)!’ (Jastrow 1978: 310); and Harran-Urfa *di-niṭni* ‘give (M.SG) me!’, *du-gul-li* ‘tell (M.SG) me!’ (see below in §3).

2.3.2. Derivation

In both groups there are nine forms for tri-consonantal roots and two forms for quadri-consonantal roots that are used for the derivation of verbs. A tenth tri-consonantal form is limited to a very few verbs. Each form has a set of patterns for the perfect and imperfect base as well as for the active and passive participle.³²

A characteristic feature of the dialect of Mardin and its surroundings is that Form II verbs, which usually have causative and intensifying functions, are frequently used instead of the corresponding Form I verbs if they govern a direct object in plural:

- (1) a. Mardin (Jastrow 1978: 164)
- | | | |
|--|------------|------------------|
| <i>yə-qṭaṭ-lu</i> | <i>ṭūd</i> | <i>mə-d-dawm</i> |
| IPF.3.M.SG-cut.FORM1-for.3.M.SG | stick.SG | from-DEF-tree |
| ‘He cuts himself a stick from the tree.’ | | |
- b. *yə-qatṭəf-lu* *ṭūd-ayn* *mə-d-dawm*
 IPF.3.M.SG-cut.FORM2-for.3.M.SG stick.PL from-DEF-tree
 ‘He cuts himself sticks from the tree.’

³² Cf. Jastrow (1978: 146–195); Lahdo (2010: 124–147); Procházka (2003: 81–82).

2.1.3. Verbal modifiers³³

The general present tense is not marked in either group, with the exception of some varieties of the Kōsa subgroup of *qəltu* Arabic.³⁴ Though Mardin Arabic does not use a modifier for the progressive present, most other *qəltu* dialects of the region do.

In the *qəltu* dialects the verbal modifier of the progressive present is *kū-*, as in *kū-yəšrab* ‘he is drinking’.

- (2) Kəndərīb (Jastrow 1999: 46)

<i>kū-tə-qṣāṣ-ūn</i>	<i>ṣīyād-ana</i>	<i>kwā t-tāžīye</i>	<i>maṣ-i</i>
MOD-IPF.2-see-PL	hunter-COP.1.SG	FOC	DEF-greyhound with-1.SG
‘You see that I am a hunter having the greyhound with me.’			

In Harran-Urfa Arabic the invariable participle *ḡaṣad*, sometimes shortened to *ḡaṣ* (< OA *qāṣid* ‘sitting’), is used.

- (3) Harran-Urfa (own data)

<i>ḡaṣad</i>	<i>a-ṣrab</i>	<i>ḡhawa</i>
MOD	IPF.1.SG-drink	coffee
‘I am drinking coffee.’		

In the *qəltu* dialects the future tense is indicated by particles and modifiers which go back to the OA conjunction *hattā* ‘so that, in order that’ (Jastrow 1978: 302). Thus in Mardin one hears *ta-nkayyaf* ‘we shall have fun’ and *ta-yəġi* ‘he will come’, in the dialects around Mardin *ta-nəšrab* ‘we shall drink’, and in Siirt/Daragözü *də-yāxev* ‘he will take’.³⁵ Imminent future tense can be expressed by a combination of *kū-* and *tə-*.

- (4) Kəndərīb (Jastrow 1999: 47)

<i>ən-nawb</i>	<i>īda</i>	<i>hawnak-we</i>	<i>kū-t-ī-rōḥ</i>
then	if	there-COP.3.M.SG	MOD-MOD-IPF.3.M.SG-go
<i>y-fəkk-a</i>			

IPF.3.M.SG-untie-3.F.SG

‘If he then comes there he will immediately untie it.’

In Harran-Urfa Arabic the future tense is expressed by the auxiliary verb *yrīd* ‘to want’: e. g. *?arīd aṣrab* ‘I will drink’ (and ‘I want to drink.’).

³³ A very good overview of the complex verbal system is Jastrow (1999), which deals with the dialect of the village of Kəndərīb. But most of Jastrow’s findings are true for the majority of the *qəltu* dialects of Anatolia.

³⁴ Jastrow (1978: 300), *bə-nqūl* ‘we say’. Jastrow denies any influence from the analogous modifier that is found in Syrian Arabic.

³⁵ In the village of Tuzlagözü *lē-* (< *li-/la-* ‘in order to’) is used: e. g. *lēbqa* ‘he will stay’ (Jastrow 1978: 303).

Continuous and habitual past tense is expressed in Mardin town by uninflected *kān* (literally ‘he was’) followed by an imperfect: e. g. *kān yəšrabūn* ‘they used to drink’. In other *qəltu* dialects (rural Mardin, Siirt) the modifier *kā-* ~ *ka-* is used together with a finite verb in the imperfect form: e. g. *kā-yəbki* ‘he was crying’.

- (5) Kəndərīb (Jastrow 1999: 47)

kān fī wēħəd m-Mədyād, ka-y-say
be.PF.3.M.SG there.is one from-Midyat MOD-IPF.3.M.SG.make
səħər
magic

‘There was one from Midyat who used to practice magic.’

In Harran-Urfa Arabic the continuous and habitual past tense is expressed by the complex predicate *čān* plus the verb in the imperfect.

- (6) Harran-Urfa (own data)

čin-na n-ākul čime
be-PF.1.PL IPF.1.PL-eat truffle.COLL
‘We used to eat truffles.’

The *qəltu* dialects usually express the present perfect by invariable modifiers such as *kəl-/kə-* followed by a verb in the perfect. Thus in Mardin Arabic (and other dialects) one finds *kəl-qəfdat* ‘she has sat down’ (Jastrow 1978: 307). Some dialects (Āzəx, rural Siirt) use *kū-* with the perfect, others (Mħallami/Kōsa) use *kūt-*.

- (7) Kəndərīb (Jastrow 1999: 50)

walla əl-bənt mō-d-bayyən kūt-harab-ət
truly DEF-girl NEG-IPF.3.F.SG-be.seen MOD-run.off-PF.3.F.SG

‘The girl is not to be seen, she has run off.’

In Harran-Urfa Arabic the present perfect is usually expressed by an active participle (cf. Procházka and Batan 2016).

- (8) Harran-Urfa (Procházka and Batan 2016: 458)

ʕAli mitɻallim ɻala sōg at-ṭaqṣi.
Ali learn.AP.M.SG on driving DEF-car
‘Ali has learned to drive a car (and now knows how).’

2.4. Syntax of basic clauses

2.4.1 Copula

A conspicuous feature of all Anatolian *qəltu* dialects is the consistent use of a copula that links the subject and predicate of a sentence without a verb. The “Anatolian copula” consists of enclitic variants of the personal pronoun and is attached to the predicate.³⁶

- (9) Mardin (Jastrow 1978: 27)

bayt-i gbīr-we
house-1.SG big-COP.3.M.SG
'My house is big.'

- (10) Mardin (Grigore 2007: 289)

mērdīn fōq əğ-ğabal-ye
PROPER.NAME on DEF-mountain-COP.3.F.SG
'Mardin is on the mountain (top).'

- (11) Mardin (Grigore 2007: 288)

ana Mērdīni-ana
1.SG inhabitant.of.Mardin-COP.1.SG
'I am an inhabitant of Mardin.'

The dialect of Siirt is an exception in that in propositions the copula is placed between subject and predicate (ex. 12). In questions, however, Siirt Arabic conforms to the other varieties (ex. 13).

- (12) Siirt (Jastrow 1978: 132)

āvi l-bənt īye malīħ-a
DEM.F.SG DEF-gril COP.3.F.SG good-F
'This girl is good.'

- (13) Siirt (Jastrow 1978: 132)

?əxt-ok malīħ-a-ye?
sister-2.M.SG good-F-COP.3.F.SG
'Is your sister good?'

2.4.2 Nominal attribution

In both groups, the default case of nominal attribution is by asyndetic linkage of the two (or more) elements. In this case the head noun never gets the definite article, and feminine heads appear in a special construct state marked by *-t*.

³⁶ The complete paradigms of three dialects are given in Jastrow (2006: 91).

The grade of determination depends on whether the last element of the phrase is determined or not (ex. 14). All *qəltu* dialects possess a prepositional marker for syndetic attribution.³⁷ It is, however, rare, being used mainly to avoid the repetition of the noun (ex. 15).

- (14) Harran Urfa (Procházka 2013: 209/3)

<i>miğann-t</i> [<i>miğann-e</i>]	<i>aż-zğār</i>
cemetery-F.CONSTR.ST	DEF-SMALL.PL
'the cemetery of the small (i. e. children)'	

- (15) Qarṭmīn (Jastrow 1978: 125)

<i>əl-faras</i>	<i>dīl-i-ye</i>
DEF-horse	GEN.MARKER-1.SG-COP.3.F.SG
'This horse is mine.' ³⁸	

Only in the remote dialects of the Kozluk-Sason region and in Hasköy is analytical nominal attribution frequent: it may even be regarded as the default case (Jastrow 1978: 126; Eksell 2006: 84).

- (16) Hasköy (Talay 2001: 79)

<i>səfēb</i>	<i>lē</i>	<i>baqqāl</i>
owner	GEN.MARKER	shop
'the owner of the shop'		

- (17) Hasköy (Talay 2001: 79)

<i>rēhān</i>	<i>lē</i>	<i>mōy</i>
mill.PL	GEN.MARKER	water
'water-mills'		

In Harran-Urfa Arabic the prepositional marker for substantive or pronominal attributes is very infrequent. Even for foreign words attribution is asyndetic: e. g. *pīkab-i* 'my pick-up truck'.³⁹

As for adjectival attributes, in the *qəltu* dialects definite noun-adjective phrases do not differ from definite noun-substantive phrases (Grigore and Bițună 2012: 553). This implies that the head noun of the phrase appears in the construct state, and that the definite article is attached to the following adjective only (whereas in

³⁷ Most forms can be traced back to **dī-la*, the etymology of which is, however, not undisputed. Jastrow (1978: 125) points out that it resembles the OA relative pronoun *alladī*. But Talay (2001: 78), fn. 23 suggests an Aramaic origin. In the former Jewish dialect of Siverek the typically Syrian form *tabaʃ* was used: e. g. *mə-nəstahī mən dīn tabaʃna* 'We are not afraid of our religion.' (Nevo 1999: 71).

³⁸ Instead of **əl-faras farasi-ye*.

³⁹ Rarely *hnīt* 'thing' is attested to express a kind of belonging, e. g., *ʔiħna*, *ʔhnīt as-sūg mā nāxudhe* 'We do not buy the one from the market'.

Arabic usually both components take the definite article). In Harran-Urfa Arabic, the article is usually attached to both components, though with the peculiarity that feminine nouns appear in the construct state.

- (18) Mardin (Jastrow 1978: 124)

<i>bənt</i>	<i>əl-əkwayys-e</i>
girl	DEF-beautiful-F
	'the beautiful girl'

- (19) Mardin (Grigore 2007: 216)

<i>bəqč-ət</i>	<i>əl-wəsx-a</i>
bale-F.CONSTR.ST	DEF-dirty-F
	'the dirty bale'

- (20) Harran-Urfa (own data)

al-ṣarab *awwali* *balāy* *al-ghaw-t*

DEF-Arab.COLL formerly without DEF-coffee-F.CONSTR.ST

al-murr-a *mā-y-ṣīr*

DEF-bitter-F NEG-IPF.3.M.SG-become

'In former times, among the Arabs, nothing happened without bitter coffee.'

In Harran-Urfa Arabic an adjectival attribute is nearly always linked to its head noun by the enclitic suffix *-in* if the head noun is indefinite⁴⁰ (exs. 21–22). The same morpheme is also used to link qualifying components other than adjectives to indefinite nouns (ex. 23). (For relative clauses see below, section 2.5.1).

- (21) Harran-Urfa (own data)

<i>ṣaḡīy-in</i>	<i>ṣaḡīr</i>	<i>ṣəmr-u</i>	<i>ṣahar</i>	<i>ṣahr-ēn</i>
child-LINKER	small	age-3.M.SG	month.SG	month.DUAL
'a small child whose age is one or two months'				

- (22) Harran-Urfa (own data)

ḥōš-in *čibīr*

house-LINKER big

'a big house'

- (23) Harran-Urfa (own data)

rabiṭ-in *mitl-ak*

friend-LINKER like-2.M.SG

'a friend like you'

⁴⁰ This suffix is also called a *tanwīn* after the term for the indefinite marker in OA.

2.4.3 Negation

Anatolian *qəltu* dialects use two different particles to negate perfect and imperfect, *mā* and *mō*. The latter also negates sentences which do not contain a verb.

- (24) Mardin (Grigore 2007: 169)

<i>mō</i>	<i>tə-ği</i>
NEG	IPF.2.M.SG-come
'You don't come.'	

- (25) Mardin (Grigore 2007: 169)

<i>mā</i>	<i>daxal</i>
NEG	enter.PF.3.M.SG
'He did not enter.'	

- (26) Mardin (Jastrow 1978: 30)⁴¹

<i>mō</i>	<i>hawn-ana</i>
NEG	here-COP.1.SG
'I am not here.'	

In Harran-Urfa Arabic, the particle *mā* is used to negate both perfect and imperfect verbs. Non-verbal predicates are negated by a combination of *m(ñ)-* and the personal pronouns (exs. 29–30).

- (27) Harran-Urfa (own data)

<i>?awwal</i>	<i>marra</i>	<i>mā</i>	<i>ligā-ha</i>
first	time	NEG	find.PF.3.M.SG-3.F.SG
'At the first time he did not find her.'			

- (28) Harran-Urfa (own data)

<i>mā</i>	<i>y-āxd-ūn-hum</i>	<i>fa-l-garāye</i>
NEG	IPF.3.M-take-PL-3.M.PL	to-DEF-village.PL
'They don't take them to the villages.'		

- (29) Harran-Urfa (own data)

<i>m-āni</i>	<i>čaddāb</i>
NEG-1.SG	liar
'I am not a liar.'	

- (30) Harran-Urfa (own data)

<i>mi-hin</i>	<i>šēn-āt</i>
NEG-3.F.PL	bad-PL.F
'They (F.PL) are not bad.'	

⁴¹ In some Anatolian dialects the enclitic pronoun is attached to the negator: e. g. Mhalami *mana hawn* 'I am not here' (Jastrow 1978).

In both groups the particle *lā* negates the imperative and the optative: e. g. Mardin: *lā yəġi* ‘He may not come!’ (Jastrow 2006: 92); Harran-Urfa: *lā tibčīn!* ‘Don’t (F.SG) cry!’

2.4.4 Word order

According to the in-depth study by Dahlgren (1998), in Anatolian *qəltu* Arabic there is a strong tendency towards SV in both dialogues and description (Dahlgren 1998: 194, 205). Unless topicalized, nominal direct objects come after the verb, as in examples (46) or (71) below.

2.5. Syntax of clause linkage

In both groups, coordinative clauses are often linked by conjunctions such as (Mardin) *w* ‘and’, *ʔamma/lākən* ‘but’, *vēya* ‘or’. However, unlinked sequences of clauses are equally frequent.

Subordinate clauses which are dependent on verbs of speaking, belief, and the like are very often not introduced by a conjunction (cf. also Lahdo 2010: 173). An example of syndetic linkage is (31).

- (31) Mardin (Grigore 2007: 261)

<i>y-qūl</i>	<i>kt⁴²</i>	<i>nəħne</i>	<i>sədqān</i>
IPF.3.M.SG-say	CONJ.that	1.PL	friend.PL
‘He said that we are friends.’			

Both syndetic (exs. 32–33) and asyndetic (ex. 34) linkage is frequently used for final and causal clauses. Final clauses are often introduced by the proclitic conjunction *ta-*.

- (32) Mardin (Grigore 2007: 313)

<i>Maryam</i>	<i>rāh-ət</i>	<i>ta-tə-ra</i>	<i>bənt-a</i>
PROPER.NAME	go-PF.3.F.SG	CONJ-IPF.3.F.SG-see	daughter-3.F.SG
‘Maryam went to see her daughter.’			

- (33) Harran-Urfa (Procházka 2010: 132/19)

<i>i-taqassal</i>	<i>ta-mā</i>	<i>y-đall</i>	<i>bī</i>	<i>wahl-u</i>
IPF.3.M.SG-be.washed	CONJ-NEG	IPF.3.M.SG-stay	in.3.M.SG	mud-3.M.SG
‘It is washed in order that no mud remains on it.’				

⁴² This conjunction is a loan from Turkish, but ultimately of Persian origin.

- (34) Tillo (Lahdo 2010: 175)

<i>kəg̃-ğ̃l-na</i>	<i>nə-tlab</i>	<i>bənt-ak</i>
MOD-COME-PF.1.PL	IPF.1.PL-ask	daughter-2.M.SG
'We came to ask for your daughter's hand.'		

In Mardin Arabic, causal clauses are often introduced by the conjunction *čənki* (< Kurdish *çunki/çimki* or Turkish *çünkü*), in Harran-Urfa Arabic by *čaman*.

- (35) Mardin (Grigore 2007: 313)

<i>yəlzəm</i>	<i>a-ṛōḥ</i>	<i>a-štari</i>	<i>akəl</i>	<i>čənki</i>
necessary	IPF.1.SG-go	IPF.1.SG-buy	food	CONJ.because
<i>ğūlān-ana</i>				
'I have to buy food because I'm hungry.'				

- (36) Harran-Urfa (own data)

<i>al-qanam</i>	<i>şār-at</i>	<i>zihīd-āt</i>	<i>hēne</i>
DEF-sheep.COLL	become-PF.3.F.SG	few-PL.F	here
<i>čaman</i>	<i>mā dall</i>	<i>muṭrah</i>	<i>yi-srah-ūn</i>
CONJ.because	NEG stay.PF.3.M.SG	place	IPF.3-graze-M.PL
'The (number of) sheep has decreased here, because there is no space left to graze (them).'			

Conjunctions are more or less constantly used in temporal clauses as well as in conditionals.⁴³

- (37) Mardin (Grigore 2007: 316)

<i>lamān</i>	<i>nəzəl-na</i>	<i>mə-l-bayt</i>	<i>kān</i>
CONJ.when	go.out-PF.1.PL	from-DEF-house	be.PF.3.M.SG
<i>yə-ği</i>	<i>salğ</i>		
IPF.3.M.SG.come	snow	'When we left the house it was snowing.'	

- (38) Mardin (Jastrow 1981: 4/3)

<i>čagāb</i>	<i>lamān</i>	<i>a-mūt</i>	<i>w</i>	<i>yə-fłəs</i>
PARTICLE	CONJ.when	IPF.1.SG-die	and	IPF.3.M.SG-go.bankrupt
<i>Paš</i>	<i>ta-yşēr</i>		<i>dūrum-u?</i>	
Q.what	MOD-IPF.3.M.SG-become	situation-3.M.SG	'When I die and he goes bankrupt, what will become of him?'	

⁴³ The conditional sentences of Mardin Arabic are described in detail by Grigore (2007: 293–308) and Grigore (2008); for the dialect of Kəndərīb see Jastrow (1999: 51–53).

- (39) Tillo (Lahdo 2010: 173)

waxt lay *talāh-t* *mə-l-karm* *aʃla* *kān*
 CONJ.when leave-PF.2.M.SG from-DEF-vineyard God be.PF.3.M.SG
ʃala *rās-ak*
 on head-2.M.SG
 ‘When you left the vineyard God was watching you.’

- (40) Mardin (Jastrow 1981: 8/18)

?awwal mā *ra-w-hu* *mən* *bʃid* *yāho* *kalb* *əbən*
 CONJ.as_soon_as see-PF.3.PL-3.M.SG from far hey dog son
kalb
 dog

‘As soon as they saw him (they said to him), “Hey dog, son of a dog!”’

- (41) Mardin (Jastrow 1981: 10/28)

abū-y *kəl-hallaf-ni* *qabəl la* *y-mūt*
 father-1.SG MOD-let.swear.PF.3.M.SG-1.SG CONJ.before IPF.3.M.SG-die
 ‘My father let me swear before he died.’

- (42) Harran-Urfa (Procházka 2010: 129/5)

yōm-in *y-ṣīr* *muṭar* *y-gūl-ūn*
 CONJ.when IPF.3.M.SG-become rain IPF.3-say-M.PL
 ‘When it rains they say ...’

- (43) Harran-Urfa (own data)

ʃugub-ma *mayyit* *abū-ha* *dāll-it-ill-u*
 CONJ.after die.AP.M.SG father-3.F.SG remain.AP-F.SG-for-3.M.SG
ibnayye
 daughter

‘After her father had passed away, he left behind a daughter.’

- (44) Harran-Urfa (own data)

ʃugub-ma *y-ṣīr* *ğubun* *ʃādēn* *ši-t-sawwi?*
 CONJ.after IPF.3.M.SG-become cheese PARTICLE Q.what-IPF.3.F.SG-do
 ‘After it has become cheese, what is she doing?’

- (45) Harran-Urfa (own data)

gabəl-ma *ti-ğī* *l-mayye* *čin-na* *ni-zraʃ*
 CONJ.before IPF.3.F.SG-come DEF-water be-PF.1.PL IPF.1.PL-plough
ʃala *l-fiddān*
 on DEF-OX

‘Before the water (*viz.* irrigation) came, we used to plough with oxen.’

- (46) Harran-Urfa (own data)

šnōn-ma *t-gūm* *ʔt-qassil* *bēt-he*
 CONJ.as.soon.as IPF.3.F.SG-get.up IPF.3.F.SG-wash house-3.F.SG
 ‘As soon as she gets up, she cleans her house.’

2.5.1 Relative clauses

There are only slight differences between *qəltu* and Harran-Urfa Arabic in the general rules regarding relative clauses. For the most parts, such clauses follow their head. In the Anatolian *qəltu* dialects these clauses are asyndetically linked to the head noun if this is indefinite. In Harran-Urfa Arabic the relative clause as a qualifying element is connected to the head noun by the linking suffix *-in* (see above, section 2.4.2).

- (47) Kenderib (Jastrow 1978: 124)

fī *ʕalay-u* *kətābe* *mō* *tə-nqəri*
 there.is on-3.M.SG inscription NEG IPF.3.F.SG-be.read
 ‘On it is an inscription that is not readable.’

- (48) Mardin (Grigore 2007: 309)

?ana *?a-dawwər* *ʕala* *kamyōn*
 1.SG IPF.1.SG-seek on lorry
y-ṭeq *yə-hməl* *ṭāt* *tōn-āt* *qərēmīd*
 IPF.3.M.SG-can IPF.3.M.SG-carry three ton-PL tile.PL
 ‘I am looking for a lorry that can carry three tons of roof tiles.’

- (49) Harran-Urfa (own data)

hināk *šī* *muṭrah-in* *yi-ṭlaf* *yi-srah*
 there there.is place-LINKER IPF.3.M.SG-go.out IPF.3.M.SG-graze
bī-hin
 in-3.F.PL
 ‘There is space where he (can) go and graze them.’

If the head of the relative clause is definite, it is linked to it by the invariable relative pronoun *la/lə/lē/lay* in the *qəltu* dialects and *al* in Harran-Urfa Arabic.⁴⁴ Both dialect groups exhibit two salient features which distinguish them from virtually all other Arabic dialects (Retsö 2004: 266 and Retsö 2009: 21): the use of the construct state with feminine nouns and the non-occurrence of the definite article on the semantically definite head noun. The latter feature is very common

⁴⁴ In Harran-Urfa Arabic the relative is identical with the definite article and thus subject to assimilation to following dentals and sibilants (see ex. 53).

in the *qəltu* dialects,⁴⁵ but only occasionally found in Harran-Urfa Arabic. Thus for the Anatolian *qəltu* dialects there are completely identical structures for (definite) noun-substantive, noun-adjective, and noun-clause attributions.

- (50) Mardin (Grigore 2007: 309)

<i>ṣarab-at</i>	<i>la</i>	<i>xalf-na</i>	<i>t-rīd</i>
car-F.CONSTR.ST	REL	behind-1.PL	IPF.3.F.SG-want
<i>ta-t-fūt-na</i>			
CONJ-IPF.3.F.SG-pass-1.PL			

'The car behind us wants to pass us.'

- (51) Mardin (Grigore 2007: 310)

<i>kalb</i>	<i>la</i>	<i>gabbar-tu</i>	<i>fə-hawš-na</i>
dog	REL	raise-PF.1.SG	in-court-1.PL

'the dog which I have raised in our court'

- (52) Qarṭmīn (Jastrow 1978: 124)

<i>ayn-i</i>	<i>ḥənt-ət</i>	<i>la-tə-yə-hīt-k</i>
Q.where-COP.3.F.SG	wheat-F.CONSTR.ST	REL-FUT-IPF.3.M.SG-give-2.M.SG

'Where is the wheat which he will give to you?'

- (53) Harran-Urfa (Procházka 2003: 83)

<i>zlim-t</i>	<i>as</i>	<i>sōlaf-t-u</i>	<i>minu?</i>
man-CONSTR.ST ⁴⁶	REL	speak-PF.2.M.SG-3.M.SG	Q.who

'Who is the man you spoke with?'

- (54) Harran-Urfa (own data)

<i>al-ḥunt-it</i>	<i>al</i>	<i>mā</i>	<i>t-tiṣawwal</i>
DEF-wheat-F.CONSTR.ST	REL	NEG	IPF.3.F.SG-be.soaked
<i>y-ṣīr</i>			
IPF.3.M.SG-become			

xubuz-he

mitīn

bread-3.F.SG hard

'Bread from wheat which has not been soaked becomes hard.'

- (55) Harran-Urfa (own data)

<i>al-ğild</i>	<i>al</i>	<i>y-ħuṭṭ-ūn</i>	<i>bī</i>	<i>as-samin</i>
DEF-skin	REL	IPF.3-put-M.PL	in.3.M.SG	DEF-ghee

'the skin in which they put the ghee'

There is evidence of relative clauses in Harran-Urfa Arabic which contain a head noun carrying the suffix *-in* although it is semantically clearly definite.

⁴⁵ Examples showing the definite article on the head noun are also attested, e. g. from Tillo *aq əl-kurmanč lay kəg-ğaw mən qabəl* 'these Kurds that came earlier' (Lahdo 2010: 178).

⁴⁶ The word *zlime* 'man' has the morphological feminine ending but is, of course, masculine by natural gender.

- (56) Harran-Urfa (own data)

ghawt-in ištarē-nā-ha
coffee.IDF-LINKER buy-PF.1.PL-3.F.SG
'the coffee which we have bought'

- (57) Harran-Urfa (own data)

?awwal yōm-in ġīl-na
first day.IDF-LINKER come-PF.1.PL
'on the first day we were here'

In clauses without a verb, the relative pronoun merges with the copula in *qəltu* Arabic. In Harran-Urfa Arabic again the suffix *-in* is used to link the two elements:

- (58) Qarṭmīn (Jastrow 1978: 138)

tə-yə-ṛa la-wwe fa-l-ard
FUT-IPF.3.M.SG-see REL-COP.3.M.SG on-DEF-ground
'He will see the one who is sitting on the ground.'

- (59) Harran-Urfa (own data)

sint-in intē ṣāyir
year-LINKER 2.M.SG become.AP.M.SG
'in the year you were born'

3. Additional notes on contact-phenomena

All Arabic dialects spoken in eastern Anatolia exhibit features that indicate the influence of the two major languages of the region, Turkish and Kurdish. Neo-Aramaic may have played a more significant role in the past, particularly in the region adjacent to Tur Abdin. But it must be emphasized that the local distribution of contact-phenomena is highly variable. As was pointed out in Procházka (2006–2007: 116), the bulk of idiosyncratic dialectal features is found in "language islands" (especially in the Kozluk-Sason group). On the whole, the Bedouin-type dialects of the Harran-Urfa region show much less Turkish (and Kurdish) influence than the *qəltu* dialects.⁴⁷

Most striking is the Turkish and Kurdish impact on the lexicon of the region's Arabic, resulting in a large number of loanwords that constitute up to one-third of the vocabulary.⁴⁸ A clear indicator of intensive contact or even bilingualism are the

⁴⁷ For possible reasons cf. Procházka (2006–2007: 115–117).

⁴⁸ For the village of Daragözü, see Vocke and Waldner (1982: xlivi). Exhaustive lists of the Turkish and Kurdish loanwords in the dialect of the village of Tillo are found in Lahdo (2010: 203–214, 217–219).

numerous calques found in phraseology. Frequently found⁴⁹ is the copying of the so-called phrasal verbs of Turkish and Kurdish, which consist of the verb ‘to do, to make’ (Turk. *etmek*, Kurd. *kirin*) with a following noun. Examples are *sawa yārdəm* (cf. Turkish *yardım etmek*) ‘to help (lit. to make help)’, and *sawa ġāmērtīye* (cf. Kurdish *camērtî kirin*) ‘to act generously (lit. to make generosity)’ (Talay 2006–2007: 184). Examples of idiomatic calques: from Harran-Urfa *?imil ġēnak!* ‘Delight your eyes!’ which is a calque of the Turkish *gözü doldur*, literally ‘Fill the eye!'; and from Tillo *bəqīna fə-n-nəşṣ* ‘we remained helpless’, which is a calque of the Turkish *ortasında kaldık*, literally ‘we stayed in the middle’ (Lahdo 2010: 174).

A consequence of the extensive adoption of Turkish and Kurdish vocabulary is that all dialects spoken in eastern Anatolia have phonemes alien to Old Arabic. Particularly conspicuous are /p/, /ž/ and /v/ (the latter two are not found in Harran-Urfa), which overwhelmingly occur in loanwords. The situation is different for /č/ and /g/: in the *qəltu* dialects these two phonemes are also more or less restricted to loanwords,⁵⁰ whereas in Harran-Urfa Arabic they are the regular reflexes of OA /k/ and /q/ when in certain vowel settings.

In the Kozluk-Sason group the velarized variants of post-dentals and sibilants have, under the influence of Kurdish, lost their phonemic status (Talay 2006–2007: 181).⁵¹

Morphological changes due to the influence of Turkish and Kurdish are not numerous. In the isolated Kozluk-Sason group the system of determination and indetermination has been remodeled according to Kurdish (and less so Turkish) patterns. This explains the frequent omission of the definite article, especially in subject position.⁵² The absence of determination marking can be explained by the fact that both adstratum languages do not have a definite article. But on the other hand these dialects exhibit the enclitic indefinite marker *-ma* which is used exactly like the Kurdish indefinite marker *-ek* that is likewise enclitic.

- (60) Mutki (Akkuş 2016: 39)

ifī atsūra-ma fo fəstox
there.is bird-IDF on roof
'There is a bird on the roof.'

In the same dialect group, inversion of the unit and decimal positions in the numerals from 11 to 19 has taken place. Thus numerals such as *fašra w səlāse* ‘13’ and

⁴⁹ This is particularly true of the *qəltu* dialects, but less common in Urfa Arabic.

⁵⁰ In a few roots *g* is also found in Arabic words as a consequence of the generalization of assimilated consonants: *gəbər – yəgbaṛ* ‘to grow’. Cf. Jastrow (1978: 47) for further examples.

⁵¹ For further phonological contact phenomena see Talay (2006–2007: 181–183), and Lahdo (2010: 191–192).

⁵² Cf. Jastrow (1973: 91), Talay (2006–2007: 185).

fašra sätte ‘16’ follow the Turkish model (*on iç*, *on altı*). Talay (2001: 78, fn. 26) reports that in the village of Hasköy nominal attribution is sometimes done by using the Kurdish linking suffix *-ē*: e. g. *šəllī-yē rabīh* ‘spring rain’.

Probably the most striking contact phenomenon on the syntactical level is the consistent use of a copula in all the *qalṭu* dialects (see above, section 2.4.1). As this feature is found in Kurdish, Turkish, and Aramaic, but not in other Arabic dialects, it can certainly be traced back to the influence of these adstratum languages.

As is pointed out by Talay (2006–2007: 183) and Jastrow (2011: 91), there are many similarities between the tense and aspect categories expressed by verbal modifiers in the *qalṭu* dialects (see above, section 2.3.3) and the Turkish verbal system. The existence of similar categories in many other Arabic dialects clearly indicates that they are not a direct product of Turkish influence, but the degree of coincidence is so striking that one can assume that the Arabic verbal system has developed at least some of these categories by analogy with the Turkish model. This may be especially true for constructions like *ka-tə-taftah* ‘you would have opened’, which is rather unusual elsewhere⁵³ and morphologically a copy of the Turkish suffix *-ecek-ti*, a combination of future and perfect tense markers like the Arabic *ka-tə-*.

Equally influenced by Turkish is the use in Harran-Urfa Arabic of the participle to express evidentiality (i. e. like the Turkish *mış*-perfect). This is very often found in narrative speech when the speaker refers to an event which was not personally observed or experienced.⁵⁴ For example, in a story about a girl who went mad, the narrator frequently used participles instead of perfect forms:

- (61) Harran-Urfa (Procházka and Batan 2016: 464)

<i>gāyil</i>	<i>Šēx Məṭar</i>	<i>ʔənṭ-ū-ni</i>	<i>hāt-ū-li</i>
say.AP.M.SG	PROPER.NAME	give-IMP.M.PL-1.SG	give-IMP.M.PL-to.1.SG
<i>hdūm-ha!</i>	<i>ʔmnawwš-īn</i>	<i>ʔhdūm-ha</i>	...
clothes-3.F.SG	hand.AP-M.PL	clothes-3.F.SG	
<i>minṭī-he</i>	<i>gāym-e</i>	<i>ti-lbas</i>	<i>gāyl-e</i>
give.AP.M.SG-3.F.SG	stand.up.AP-F.SG	IPF.3.F.SG-put.on	say.AP-F.SG
‘Sheikh Məṭar said, “Give me her clothes!” They handed him her clothes; and when he gave them to her, she started to put them on and said ...’			

Another possible influence of Turkish in Tillo Arabic is that temporal clauses usually precede the main clause (Lahdo 2010: 174). Grigore (2007: 168) claims that the particle *də-*, which emphasizes an imperative (see above, §2.3.1), goes back to the Turkish interjection *haydi* ‘let’s go!’ But this is doubtful as Turkish possesses a distinct suffix to intensify imperatives (*-sana/-sene*) and the use of *haydi*

⁵³ In the dialect of the Jews of Kurdistan the same combination of prefixes is found in a very similar function: e. g. *kān t-īmūtūn* ‘they had died’ (Jastrow 1990: 66).

⁵⁴ For a more detailed analysis cf. Procházka and Batan (2016: 464–465).

together with such forms is only optional.⁵⁵ Somewhat more likely is to assume influence of the particle *de* which is widely used in Kurmanjî Kurdish in the same function (Thackston 2006: 206), e. g. *de rûne* ‘Do sit down!’⁵⁶ However, this particle may well be of Arabic origin (probably a reflex of the OA demonstrative *dā/đī*) as it is also found in various bedouin-type dialects such as Harran-Urfa Arabic and even Baghdadi Arabic (Erwin 1963: 140). Similar forms are also attested for Aleppo, e. g. *dē ɻəl-la* ‘Come on, tell it to her!’ (Sabuni 1980: 76), and Bahraini Arabic, e. g. *d-rūhi yumma, trayyagay* ‘Go (F), my dear, and have your breakfast!’ (Holes 2016: 304) for both of which Kurdish influence is less likely.

In summary: With the exception of the Kozluk-Sason group the impact of Turkish and Kurdish on the region’s dialects is surprisingly low outside of the lexicon. Kurdish has clearly had less impact on the region’s Arabic than Turkish despite its longer duration of contact and the fact that Kurdish is still the area’s dominant everyday language. A possible explanation for this could be that Kurdish has never had such a dominant role as Turkish in the media and in school education, both of which strongly influence all layers of the population.

4. Short glossed text

Urfa Arabic (from the village of Yalnızca)

Speaker: male, born ca. 1965, farmer

Recorded by Stephan Procházka in May 2011

- (62) *?awwali al-mille kull ə.byūt-hum ə.hnīt-hum ə.b=galb
formerly DEF-people all house.PL-3.M.PL stuff-3.M.PL in=heart
ağ-ğarye
DEF-village*

‘In former times the houses and the stuff of the people were all inside of the village.

- (63) *?alħaz al ɻind-u gāq gām
now REL at-3.M.SG land stand.up.PF.3.M.SG
y-sawwi b=galħeb gāq-t-u bēt
3.M.SG-make.IPF in=heart land-F.CONSTR-3.M.SG house
Now, the one who owns land has begun to make a house on his estate.*

⁵⁵ Furthermore, the Turkish *haydi* is widely attested in the Anatolian dialects as *hēdi*, *ēdi* etc. (Vocke and Waldner 1982: 451), which makes it rather unlikely that it has been shortened to become a prefix.

⁵⁶ I owe this suggestion to Geoffrey Haig who also provided the example.

- (64) *ti.dāqāt-at* *al-mille* *dahhiğ* *hīcid* *al-hagg*
 disperse.PF-3.F.SG DEF-people look.IMP.M.SG SO DEF-truth
 The people have dispersed. Look! That's the truth!
- (65) *w-allā* *?awwali* *kull-u* *qanam* *al-bēt* *at*
 by-God formerly all-3.M.SG sheep DEF-house REL
t-rūh *bī* *?arbaṣīn* *xamsīn* *īmy-it*
 2.M.SG-go.IPF in.3.M.SG 40 50 100-F.CONSTR
dābb-it *qanam* *šī*
 beast-F.CONSTR sheep there.is
 By God, formerly there were all sheep. Whatever house you entered
 there were 40, 50, 100 head of sheep.
- (66) *?alḥaz* *mā.min* *mā.min* *alḥaz* *qanam* *zihīd*
 now NEG.there.is NEG.there.is now sheep few
čīr *zihīd* *bass* *b=aḡ-ḡbāl* *šī*
 very few only in=DEF-mountain.PL there.is
 Now there aren't. There aren't. Now there are few, very few sheep, only
 in the mountains are there sheep.
- (67) *hināk* *b=aḡ-ḡbāl* *šī* *qanam* *ṣaman* *hināk* *šī*
 there in=DEF-mountain.PL there.is sheep because there there.is
mutrah-in *yi-tlaṣ* *yi-srah* *bī-hin* *?ama*
 place-LINK 3.M.SG-go.up.IPF 3.M.SG-graze.IPF in-3.F.PL but
hēne *mā.min*
 here NEG.there.is
 Over there in the mountains, there are sheep, because there is enough
 place to graze them. But here aren't any.
- (68) *mutrah* *al* *bī-he* *mayye* *?alḥaz* *mā* *dall*
 place REL in-3.F.SG water now NEG remain.PF.3.M.SG
bī-he *qanam*
 in-3.F.SG sheep
 At places where there is water, there no sheep now remain.
- (69) *kull-u* *b=aḡ-ḡbāl* *hadāk* *bāṣ-u* *ṣala*
 all-3.M.SG in=DEF-mountain.PL DEM.M.SG sell.PF-3.M.PL on
?ahil *aḡ-ḡbile*
 people DEF-mountain
 All of them are in those mountains. They have sold (them) to the mountain
 people.

- (70) *?awwali al-qanam yōm.in y-ḥalb-ūn-ha fugub.ma*
 formerly DEF-sheep when 3-milk.IPF-M.PL-3.F.SG after.CONJ
y-ḥalb-ūn-ha al-ḥurma ač-čibīr-e al b=al-bēt
 3-milk.IPF-M.PL-3.F.SG DEF-woman DEF-old-F.SG REL in=DEF-house
hiyya ḥ. t-ḥarf
 3.F.SG 3.F.SG-know.IPF

In former times, when they milked them, after they have milked them ...
 the old woman (did it), the one who was at home, she knew it.

- (71) *ḥurm-it ač-čibīr-e al-umma hadīč*
 woman-F.CONSTR DEF-old-F.SG DEF-mother DEM.F.SG
t-āxd al-ḥalīb w=ḥa=n-nār
 3.F.SG-take.IPF DEF-milk and=on=DEF-fire
t-fawwr-u ɻawwal fāl
 3.F.SG-bring.to.the.boil.IPF-3.M.SG first time
 The old woman, the “mother” is the one who takes the milk and first brings
 it to the boil on the fire.
- (72) *fugub.ma t-fawwr-u t-xallī*
 after.CONJ 3.F.SG-bring.to.the.boil.IPF-3.M.SG 3.F.SG-let.IPF.3.M.SG
t-sawwi dinlenme ta=yu-brud gutma
 3.F.SG-make.IPF rest [Turkish] PURP=3.M.SG-cool.IPF a.little
 After she has brought it to the boil she lets it rest so it cools down a little.'

Abbreviations

AP	Active participle	IDF	Indefinite
COLL	Collective	IMP	Imperative
CONJ	Conjunction	IPF	Imperfective
CONSTR.ST	Construct state	M	Masculine
COP	Copula	MOD	Modifier
DEF	Definite article	NEG	Negation
DEM	Demonstrative	PF	Perfective
DUAL	Dual	PL	Plural
F	Feminine	PURP	Purposive
FOC	Focus	Q	Question particle
FORM1, 2	First, second verbal stem	REL	Relativizer
FUT	Future	SG	Singular
GEN.MARKER	Genitive marker		

References

- Akin, Cahit, Jastrow, Otto & Shabo Talay. 2017. Nörşen – ein anatolisch-arabischer Dialekt der Sason-Muş-Gruppe. *Zeitschrift für arabische Linguistik* 65. 73–95.
- Akkuş, Faruk. 2016. The Arabic dialect of Mutki-Sason areas. In George Grigore & Babriel Bituňă (eds.), *Arabic varieties: Far and wide. Proceedings of the 11th Conference of AIDA – Bucharest, 2015*, 29–41. Bucharest: Editura Universității din București.
- Andrews, Peter Alford (ed.). 2002. *Ethnic groups in the Republic of Turkey*. Wiesbaden: Reichert.
- Arnold, Werner & Peter Behnstedt. 1993. *Arabisch-Aramäische Sprachbeziehungen im Qalamūn (Syrien)*. Wiesbaden: Harrassowitz.
- Benninghaus, Rüdiger. 2002. Identity in question – Some preliminary remarks on the Mḥal-lami. In Peter Alford Andrews (ed.), *Ethnic groups in the Republic of Turkey. Supplement and Index*, 179–189. Wiesbaden: Reichert.
- Dahlgren, Sven-Olof. 1998. Word order in Arabic. Göteborg: Acta Universitatis Gothoburgensis dissertation.
- Diem, Werner. 1979. Studien zur Frage des Substrats im Arabischen. *Der Islam* 56. 12–80.
- Eksell, Kerstin. 2006. Analytic Genitive. In Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics*, vol. 1, 82–85. Leiden & Boston: Brill.
- Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.
- Grigore, George & Gabriel Bituňă. 2012. Common features of North Mesopotamian Arabic dialects spoken in Turkey (Şırnak, Mardin, Siirt). In M. Nesim Doru (ed.), *Bilim düşünce ve sanatta Cizre*, 545–555. İstanbul: Mardin Artuklu Üniversitesi Yayınları.
- Grigore, George. 2007. *L'arabe parlé à Mardin: monographie d'un parler arabe «périphérique»*. Bucharest: Editura Universității din București.
- Grigore, George. 2008. Conditional structures in Mardini Arabic. *Zeitschrift für Arabische Linguistik* 49. 63–77.
- Heidemann, Stefan. 2002. Die Renaissance der Städte in Nordsyrien und Nordmesopotamien: Städtische Entwicklung und wirtschaftliche Bedingungen in ar-Raqqa und Harrān von der Zeit der beduinischen Vorherrschaft bis zu den Seldschuken. Leiden & Boston: Brill.
- Holes, Clive. 2016. *Dialect, culture, and society in Eastern Arabia*, vol. 3: Phonology, Morphology, Syntax, Style. Leiden & Boston: Brill.
- Jastrow, Otto. 1973. *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien). Grammatik und Texte*. (Erlanger Beiträge zur Sprach- und Kunstwissenschaft 46). Nürnberg: Hans Carl.
- Jastrow, Otto. 1978. Die mesopotamisch-arabischen qəltu-Dialekte, vol. 1: Phonologie und Morphologie. Wiesbaden: Steiner.
- Jastrow, Otto. 1981. Die mesopotamisch-arabischen qəltu-Dialekte, vol. 2: Volkskundliche Texte in elf Dialekten. Wiesbaden: Steiner.
- Jastrow, Otto. 1990. Der arabische Dialekt der Juden von 'Aqra und Arbīl. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1994. The qəltu dialects of Mesopotamian Arabic. In Jordi Aguadé, Federico Corriente & Marina Marugán (eds.), *Actas del congreso internacional sobre interacciones lingüísticas arabo-romances y paralelos extra-iberos*, 119–123. Zaragoza: [no publisher].

- Jastrow, Otto. 1999. Verbformen und ihre Funktionen im arabischen Dialekt von Kəndərīb (Südostanatolien). In Norbert Nebes (ed.), *Tempus und Aspekt in den semitischen Sprachen: Jenaer Kolloquium zur semitischen Sprachwissenschaft*, 43–53. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 2004. Aramäische Lehnwörter im arabischen Dialekt von Kinderib (Südost-Türkei). *Estudios de Dialectología Norteafricana y Andalusí* 8. 99–103.
- Jastrow, Otto. 2006. Anatolian Arabic. In Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics*, vol. 1, 87–96. Leiden & Boston: Brill.
- Jastrow, Otto. 2006–2007. Where do we stand in the research on the Anatolian *qəltu* dialects? *Romano-Arabica* 6–7. 63–69.
- Jastrow, Otto. 2011. Turkish and Kurdish influences in the Arabic Dialects of Anatolia. *Türk Dilleri Araştırmaları* 21(1). 83–94.
- Kern, Laurenz W. 2015. Kurden, Araber, Scheinlibanesen: Die vielschichtige Ethnizität der Mhallami. *Wiener Zeitschrift für die Kunde des Morgenlandes* 105. 189–202.
- Lahdo, Ablaħad. 2009. Some remarks on language use and Arabic dialects in eastern Turkey. *Orientalia Suecana* 58. 105–114.
- Lahdo, Ablaħad. 2010. Tillo-Arabic: A dialectological study with emphasis on social-linguistic and language-contact phenomena. Saarbrücken: Lambert Academic Publisher.
- Morony, Michael G. 2005. *Iraq after the Muslim Conquest*, 2nd ed. Piscataway, NJ: Gorgias Press.
- Nevo, Moshe. 1999. Notes on the Judaeo-Arabic dialect of Siverek. *Zeitschrift für Arabische Linguistik* 36. 66–84.
- Oppenheim, Max Freiherr von. 1939. *Die Beduinen*, vol. 1. Wiesbaden: Harrassowitz.
- Owens, Jonathan. 2006. *A linguistic history of Arabic*. Oxford: Oxford University Press.
- Procházka, Stephan & Ismail Batān. 2016. The functions of active participles in Šāwi Bedouin dialects. In George Grigore & Gabriel Bitună (eds.), *Arabic Varieties: Far and Wide. Proceedings of the 11th International Conference of AIDA – Bucharest, 2015*, 457–466. Bucharest: Editura Universității din București.
- Procházka, Stephan. 2003. The Bedouin Arabic Dialects of Urfa. In Ignacio Ferrando & Juan José Sánchez Sandoval (eds.), *AIDA, 5th Conference Proceedings. Cádiz, September 2002*, 75–88. Cádiz: Universidad de Cádiz.
- Procházka, Stephan. 2006–2007. Does geographical periphery imply linguistic periphery? The examples of the Arabic dialects of Cilicia and Urfa in southern Turkey. *Romano-Arabica* 6–7. 109–132.
- Procházka, Stephan. 2010. Genüsse aus der Steppe: Kulturgeschichtliches und Etymologisches zur Wüstentrüffel nebst einem Text im arabischen Dialekt von Urfa (Südost-Türkei). *Wiener Zeitschrift für die Kunde des Morgenlandes* 100. 119–135.
- Procházka, Stephan. 2013. Interesting facts on Ancient Mounds – Three Texts in the Bedouin Arabic Dialect of the Harran-Urfa-Region (Southeastern Turkey). In Clive Holes & Rudolf de Jong (eds.), *Arabia and its dialect heritage. Essays in honour of Bruce Ingham*, 203–213. Leiden & Boston: Brill.
- Procházka, Stephan. 2018. The Northern Fertile Crescent. In Clive Holes (ed.), *The historical dialectology of Arabic: Linguistic and sociolinguistic approaches*. Oxford: Oxford University Press [forthcoming].
- Retsö, Jan. 2004. Relative-clause marking in Arabic dialects: A preliminary survey. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 263–273. Leiden & Boston: Brill.

- Retsö, Jan. 2009. Nominal attribution in Semitic: Typology and diachrony. In Janet C. E. Watson & Jan Retsö (eds.), *Relative clauses and genitive constructions in Semitic*, 3–33. Oxford: Oxford University Press.
- Sabuni, Abdulghafur. 1980. Laut- und Formenlehre des arabischen Dialekts von Aleppo. Frankfurt: Peter Lang.
- Sasse, Hans-Jürgen. 1971. Linguistische Analyse des arabischen Dialekts der Mhallamīye in der Provinz Mardin (Südosttürkei). Munich: Ludwig-Maximilians-Universität dissertation.
- Talay, Shabo. 2001. Der arabische Dialekt von Hasköy (Dēr-Khāṣ) Ostanatolien. Teil I: Grammatikalische Skizze. *Zeitschrift für Arabische Linguistik* 40. 71–89.
- Talay, Shabo. 2006–2007. The influence of Turkish, Kurdish and other neighbouring languages on Anatolian Arabic. *Romano-Arabica* 6–7. 179–188.
- Thackston, Wheeler M. 2006. *Kurmanji Kurdish: A reference grammar with selected readings*. Unpublished Manuscript, Harvard University. https://sites.fas.harvard.edu/~iranian/Kurmanji/kurmanji_1_grammar.pdf (accessed 20 February 2018).
- Vocke, Sibylle & Wolfram Waldner. 1982. *Der Wortschatz des anatolischen Arabisch*. Erlangen: W. Waldner.
- Weninger, Stefan. 2011. Aramaic-Arabic language contact. In Stefan Weninger (ed.), *The Semitic Languages: An International Handbook*, 747–755. Berlin & Boston: De Gruyter Mouton.
- Wittrich, Michaela. 2001. *Der arabische Dialekt von Azex*. Wiesbaden: Harrassowitz.

2.5. The Neo-Aramaic dialects of eastern Anatolia and northwestern Iran

Geoffrey Khan

1. Genetic affiliation

The Neo-Aramaic dialects of Anatolia and northwestern Iran belong to two major subgroups of Neo-Aramaic. These are the Central Neo-Aramaic (CNA) subgroup and the North-Eastern Neo-Aramaic (NENA) subgroup.¹ The dialect geography of Neo-Aramaic in the region has undergone radical changes over the last one hundred years, the most stark result of which has been the displacement of all the NENA-speaking communities from Turkey (see §2). The following geographical description, therefore, relates to the situation that existed at the beginning of the twentieth century, before these major upheavals. The CNA subgroup of dialects were spoken by Christian communities in southeastern Turkey in the region of Tûr 'Abdîn, which extends from the town of Mardin in the west up to the boundary of the Tigris river in the east and north. The main component of this subgroup is the cluster of dialects of the Neo-Aramaic variety generally known in the academic literature as Turoyo.² The main dialect split in Turoyo is between the dialect of the town of Midyat and the dialects of the surrounding villages. The differences between the Turoyo dialects are small and they are mutually comprehensible (Jastrow 1985; Ritter 1990; Waltisberg 2016). In addition to the Turoyo cluster one other dialect is known to have existed in the CNA subgroup. This is the dialect of the Christians of the village of Mlahso (now Yünlüce), situated near Lice in northern Diyarbakir province, which is related to Turoyo but exhibits a number of significant differences (Jastrow 1994a, 2011).

The NENA dialects of the region were spoken to the east of the CNA area across southeastern Turkey up to Lake Urmî in West Azerbaijan province in northwestern Iran. The boundary between the areas of the two dialect subgroups is formed by the Tigris river. The NENA dialect area of the region is an extension of the NENA dialect area of northern Iraq (Khan, this volume, chapter 3.4). As in northern Iraq, the NENA dialects in the region dealt with in this chapter exhibit considerable diversity. In southeastern Turkey the vast majority of the dialects were spoken by Christian communities, with only isolated Jewish communities.

¹ The term Central Neo-Aramaic was proposed by Tsereteli (1977) for Turoyo in his classification of the Neo-Aramaic dialects. The term North-Eastern Neo-Aramaic was coined by Hoberman (1988: 557). For the basic features of NENA see chapter 3.4.

² The most commonly used term used by native speakers of Turoyo to refer to their language is Şurâyt.

In the region of Urmi in northwestern Iran, on the other hand, there were several NENA-speaking Jewish communities, in particular in the towns of Urmi, Šəno (official name Ushnuye), Solduz (official name Naghade) and Sablagh (now Mahabad) (Garbell 1965; Khan 2008), in the area of Salamas north of the Urmi plain (Duval 1883; Mutzafi 2015), and in adjacent towns that are now situated in the east of Turkey, such as Başkale and Gawar (official name Yükseкова). The Jewish dialects of this area are closely related and form a single cluster, which is referred to here as the J. Urmi cluster. These are an extension of the so-called trans-Zab cluster of Jewish dialects in Iraq (Khan, this volume, chapter 3.4; Mutzafi 2008). A few isolated NENA-speaking Jewish communities were found elsewhere in southeastern Turkey, for example in Challa (Čäl) (Fassberg 2010) in Hakkâri province and Cizre (Nakano 1973) in Şırnak province, who spoke dialects very closely related to those of the so-called *lisana deni* cluster of Jewish dialects of northwestern Iraq (Khan, this volume, chapter 3.4).

The Christian NENA dialects of the region may be classified into several clusters. These include the following:

- i. The Bohtan cluster, spoken in villages in the area that is now the Şırnak and Siirt provinces of Turkey (referred to below as C NENA Bohtan).³
- ii. The Cudi cluster, spoken in villages in the area of the Cudi mountain (Cudi Dağı) that is now in the Şırnak province of Turkey (referred to as C NENA Cudi).
- iii. The Tiyare cluster, divided into Upper Tiyare and Lower Tiyare, spoken in villages on the western side of what is now the Hakkâri province of Turkey (referred to as C NENA Tiyare).
- iv. The Txuma cluster, spoken in villages lying to the east of the Lower Tiyare area (referred to as C NENA Txuma).
- v. The Hakkari cluster, spoken in a variety of villages in the Hakkari mountains east and north-east of Tiyare, including villages in the area of lake Van, and Salamas in northwestern Iran (referred to as C NENA Hakkari).
- vi. The cluster of dialects spoken in the far east of Turkey in the areas of Šamməsdin, Gawar and in the mountains of Tergawar over the border in northwestern Iran (referred to as C NENA Šamməsdin-Gawar).
- vii. The Urmi cluster, which includes varieties of what is best considered a single dialect spoken by Christians in villages situated on the plain of Urmi and within the town of Urmi (referred to as C. Urmi).

³ Fox (2002, 2009) uses the regional term “Bohtan” to refer to the dialect specifically of the villages of Ruma, Shwata, and Borb, which are distinct from other dialects of the Bohtan cluster in some features. For the sake of clarity, examples from Fox’s data are cited here under the name of the village Ruma.

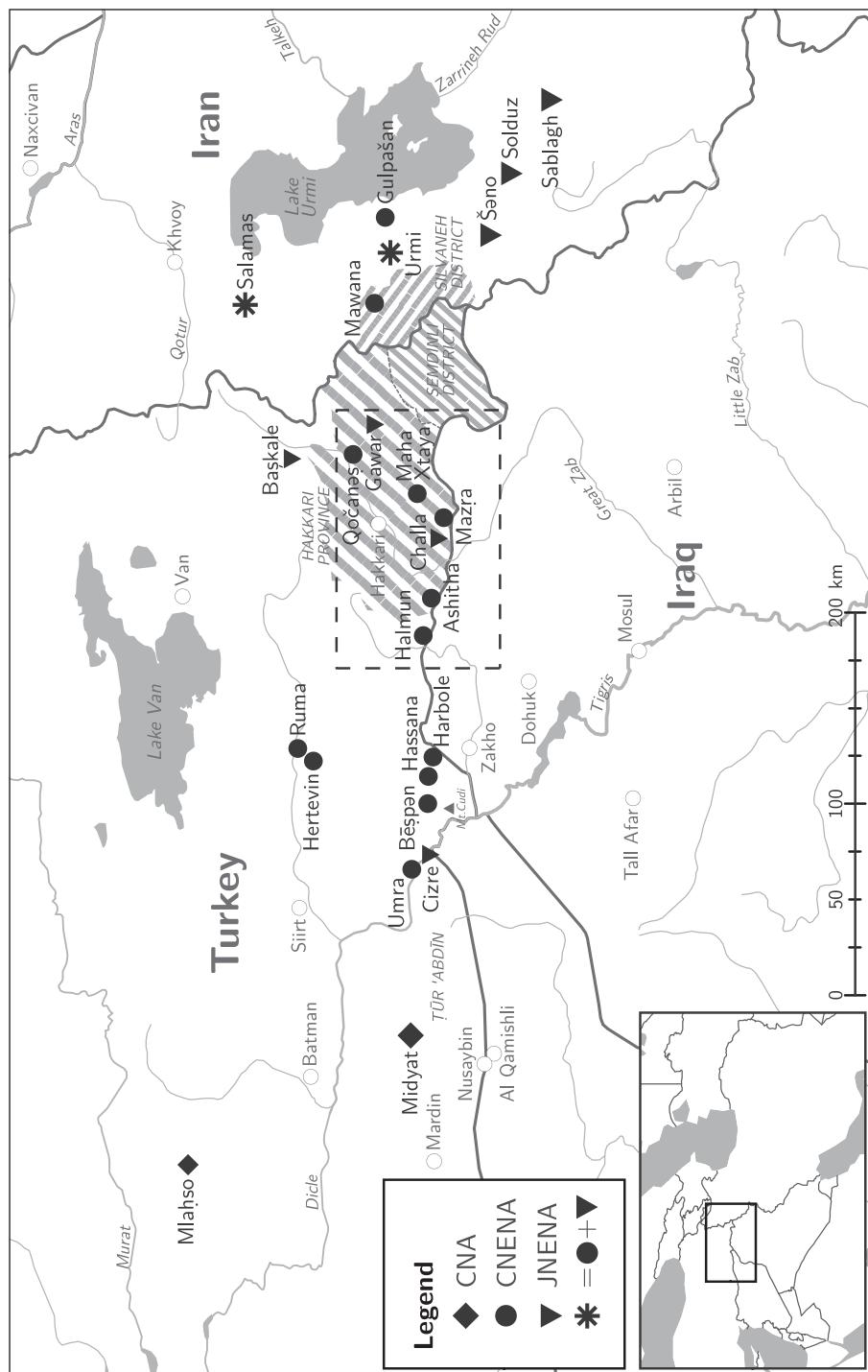


Figure 1: Location of Neo-Aramaic dialects discussed in this chapter

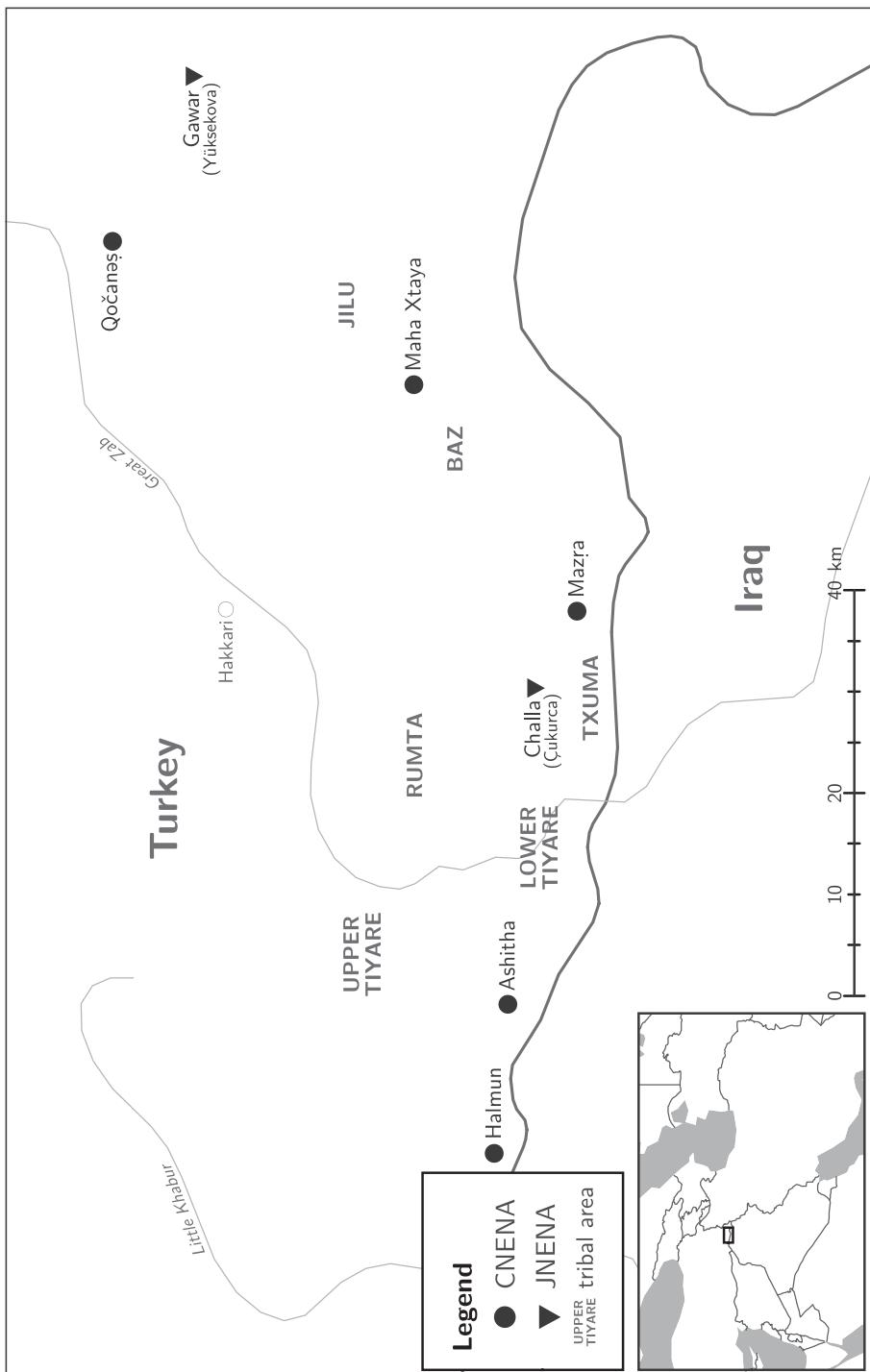


Figure 2: Christian tribal areas in the Hakkari region (southeastern Turkey)

2. History of the speech community

The geographical division between the Central Neo-Aramaic subgroup and the NENA subgroup coincides with the ancient border between the Romans and Parthians, and later between the Byzantines and the Sassanians (Kim 2008). The linguistic boundary between the two subgroups also coincides with an early Christian ecclesiastical division between the Jacobite (Syrian Orthodox) denomination of the communities west of the Tigris and the Nestorian (Church of the East) denomination of the communities to the east of the Tigris. The Christianity of the NENA-speaking communities has become more diverse in recent centuries, especially after the formation of the Chaldean Church, which is in communion with Rome, and the activities of Protestant missionaries in the 19th century. There are only sparse historical records relating to the various Neo-Aramaic-speaking settlements of the region. Some linguistic aspects of the dialects, however, give us insights into the history of the communities. The diversity of the NENA dialects, for example, can be interpreted as a reflection of the antiquity of settlement of the communities. Another factor that is likely to have had a bearing on this diversity is the fact that many of the NENA-speaking communities had the status of semi-independent tribes (*'aširatte*). As elsewhere in the NENA subgroup, the Jewish NENA dialects of the region exhibit major differences in their structure from the neighbouring Christian NENA dialects. This reflects the differing migration histories of the communities. The Jews of Urmi, for example, had settled in the town at an early period whereas the Christians of the area were almost exclusively agriculturalists living in the surrounding villages and only began to settle in the town in large numbers in the late 19th century.

The Neo-Aramaic-speaking communities underwent a major upheaval during the First World War in 1915, when they suffered massacres and mass displacement from their homes in an Ottoman-led campaign in southeastern Turkey. No accurate statistics are available for the total death toll, but it is estimated that as much as half of the Neo-Aramaic-speaking population perished, either through violence, disease or starvation, possibly amounting to around 250,000. Some of the survivors of the Christian communities subsequently returned to their homes in the Tür 'Abdīn area and the villages in the vicinity of the Cudi mountain. The NENA-speaking communities in the remainder of southeastern Turkey, however, became permanently displaced. The majority settled initially in refugee camps in Iraq, then subsequently in Iraqi towns, in particular Baghdad and Kirkuk. Some of the Christians from the Bohtan region fled northwards and found safety in the Russian empire, eventually settling in the village of Gardabani in Georgia or in Krosnodar. From 1933 to 1935 about ten thousand refugees from southeastern Turkey were settled in refugee camps in northeastern Syria, then subsequently in villages on both sides of the Khabur River. During the Kurdish uprisings in the second half of the twentieth century there were further upheavals. The villages

in the Cudi region of southeastern Turkey were destroyed in the 1980s and early 1990s and the NENA-speaking population settled in Europe, mainly in France and Belgium. Many of the speakers of Turoyo left the region, settling in the Turkish cities, especially Istanbul, or emigrating, mainly to Germany, Sweden and the USA.

The vast majority of the Jews left the region in the early 1950s and settled in the newly established State of Israel.

3. Current status of the speech community

The Turoyo-speaking community of the Tûr ‘Abdîn region is now considerably depleted. A few isolated members of the community who emigrated subsequently built houses back in the homeland, but the general trend is still an inexorable migration and now only about 2,500 still live in the original homeland, mainly in the town of Midyat. The dialect of Mlahso, which was documented by Jastrow (1994a) is now, apparently, extinct.

There are now no NENA-speaking communities in southeastern Turkey. Those who had remained in the area of the Cudi mountain finally left in the 1980s when their villages were destroyed in conflict between the Turks and the Kurds. There are still some Christians living in the Urmi region, almost exclusively now in the town of Urmi rather than the villages. A large proportion of the Christian community, however, has left the region. Some have settled in the large Iranian cities, mainly Tehran. In the 19th and early 20th centuries many Christians speakers of the dialect of Urmi and other dialects of northwestern Iran and the adjacent region moved to the Caucasus. These dialects, especially the Urmi dialect, are still spoken by communities in Georgia and Armenia, and elsewhere in the former Soviet Union (Tsereteli 1970). A large proportion of speakers have now settled in the USA, in particular in California, and Australia.

In recent times the communities who settled in the Khabur region have been under considerable pressure, in particular during the current military conflict in Syria.

After the Jews settled in Israel in the 1950s there has been a gradual decline in the number of speakers of the dialects. The dialect of Challa (Čâl) is now extinct (Fassberg 2010). There are currently very few surviving speakers of the Urmi cluster of dialects who have a good knowledge of the language. In the 19th century many families from northwestern Iran, in particular from the region of Salamas, moved to Tbilisi in Georgia, and subsequently in 1950 were settled by Stalin in Almaty in Kazakhstan, where several speakers of the dialect can be found today.

4. Sociolinguistic situation

There was a general state of multilingualism across the Neo-Aramaic speech communities of the region. In the Tûr ‘Abdîn area speakers of Turoyo often also speak vernacular Arabic (of the town of Mardin or the Mhallamî bedouin), Kurdish and Turkish, the latter being the official state and school language. In addition, Classical Syriac is used as a liturgical language and also, by some learned members of the community, as a written language. In Sweden in the 1980s an official written form of Turoyo in the Roman alphabet was created by Yusuf Ishaq and his collaborators (Ishaq 1990; Heinrichs 1990).

In the NENA-speaking area of southeastern Turkey at the beginning of twentieth century before the upheavals there was general bilingualism in Neo-Aramaic and Kurdish. There were also Armenian-speaking communities in the region. There was little exposure to Arabic, though Arabic loanwords within the dialects suggest that the contact with Arabic was greater at an earlier historical period. The communities from southeastern Turkey who settled in the Khabur area of Syria in the 1930s speak Arabic in addition to their native NENA dialects. The communities of the Khabur generally preserved their individual dialects and did not develop a koine dialect as was the case among the refugee communities in Iraq (Talay 2008a).

In northwestern Iran speakers of the NENA dialects had contact with Kurdish and Azeri Turkish. Nowadays most Christian speakers of NENA who still live in the area also speak Azeri, the vernacular of the Muslim population in this area, and also Persian, the official language of Iran, but not Kurdish. There are, however, numerous Kurdish loanwords in the NENA dialects of this area and their morphology indicates that they form an older historical layer of the lexicon than the many Azeri words. This indicates that there must have been a more widespread knowledge of Kurdish in the NENA communities at an earlier period (Khan 2016, vol. 3: 1–3). Western missionaries who were active among the Christian communities of the area in the middle of the 19th century developed a literary form of the Christian Urmi dialect written in Syriac script. This form of literary language became widely used by learned native speakers and is still used to this day (Murre-van den Berg 1999).

The NENA speakers who live in the Caucasus speak also Russian, Armenian or Georgian. In Armenia the Christian NENA speakers live in villages together with Armenians and there are many intermarriages. Many of the native Armenian-speakers in these villages also speak NENA.

The surviving Jewish speakers of NENA from northwestern Iran who settled in Israel are fluent in Modern Hebrew, which is now their primary language. The communities who have survived in Almaty now generally prefer to speak in Russian, especially the young generations.

5. Phonology

Dialects differ in the degrees of conservatism in phonology. Innovations in phoneme inventories are typically conditioned by language contact, the less conservative dialects eliminating sounds that do not occur in contact languages. In some cases language contact can facilitate archaic retention. In what follows a selection of cases will be presented from the consonant system. The term ‘earlier Aramaic’ is used to refer to what can be assumed to be the ultimate historical form of features attested in NENA.

5.1. *bgdkpt* consonants

In earlier Aramaic the stop consonants *bgdkpt* developed fricative allophones after vowels, which can be represented *[b], *[g], *[d], *[k], *[p], *[t]. In both the CNA and NENA dialects these fricative allophones became phonemicized, with the result that minimal pairs are found with stops and fricatives, e. g. CNA Turoyo *kətyo* ‘he is’ vs. *kəθyo* ‘she comes’, NENA Ashitha (C. Tiyyare) *tela* ‘fox’, *θela* ‘she came’ (Borghero 2005: 43).

Turoyo in the CNA subgroup is the most conservative dialect in this feature, since it is the only dialect in the region that has preserved all the original fricative forms of the *bgdkpt* consonants (Jastrow 1985: 6–10), e. g.

<i>táwno</i>	‘straw’	< <i>tabnā</i>
<i>ráglo</i>	‘foot’	< <i>raḡlā</i>
<i>bóxe</i>	‘he weeps’	< * <i>bāxē</i>
<i>íðo</i>	‘hand’	< * <i>īdā</i>
<i>káfno</i>	‘hunger’	< * <i>kaṭnā</i>
<i>tlóθo</i>	‘three’	< * <i>tlāṭā</i>

CNA Mlahso is less conservative in that the interdental fricatives **d* and **t* have merged with /s/ and /z/ respectively (Jastrow 1994a):

<i>tevnó</i>	‘straw’	< <i>tabnā</i>
<i>regló</i>	‘foot’	< * <i>raḡlā</i>
<i>boxé</i>	‘he weeps’	< * <i>bāxē</i>
<i>izó</i>	‘hand’	< * <i>īdā</i>
<i>nofél</i>	‘he falls’	< * <i>nāpel</i>
<i>tlosó</i>	‘three’	< * <i>tlāṭā</i>

Mlahso differs from Turoyo also in the reflex of fricative **b*. In Turoyo this reflex is /w/, which coincides with the reflex of historical **w*. In Mlahso the reflex of **b* is the labio-dental /v/, which is distinct from the reflex of historical **w*:

Turoyo	<i>táwno</i> ‘straw’ (< *tabnā)	<i>gáwzo</i> ‘nut’ (< *gawzā)
Mlaḥso	<i>tevnó</i> ‘straw’ (< *tabnā)	<i>gawzó</i> ‘nut’ (< *gawzā)

CNA, both Turoyo and Mlaḥso, is distinguished from NENA in the preservation of the fricatives /ɣ/ and /ʃ/ in the *bgdkpt* series of consonants, probably due to contact with Arabic, which has these sounds in its phoneme inventory (Jastrow 2015). In all NENA dialects of the region historical *ḡ has been lost and shifted either to a laryngeal // or to Ø. The historical development was [ḡ] > /' / > / / > Ø, in which the velar first shifted to a pharyngeal fricative before weakening to a laryngeal and zero (Tsereteli 1990). The reflex /' / is, accordingly, more archaic than the reflex Ø, e. g.

Hertevin (C NENA Bohtan)	<i>máše'</i> (< *maššeḡ; cf. Turoyo <i>mašəḡ</i>) ‘he washes’ (Jastrow 1988: 6)
C Urmī	⁺ <i>pálli</i> (< *mpalləḡ; cf. Turoyo <i>mfalág</i>) ‘divide!’

Throughout NENA the reflex of the fricative *b coincides with the reflex of *w, which is /w/ throughout most of southeastern Turkey, but /v/ in the eastern periphery of the region in northwestern Iran:

Ashitha (C Tiyare NENA)	<i>kθawa</i> ‘book’ (< *kṭābā)
C Urmī	<i>ctava</i>

A historical *p̥ has generally merged with the stop /p/, e. g.

C Urmī	<i>nápəl</i> (< *nāp̥əl; cf. Turoyo <i>nófel</i>) ‘he falls’
--------	---

Within NENA of the region, several dialects have preserved the interdental fricatives. These are mainly in the C Lower Tiyare and C Txuma clusters, e. g.

Ashitha (C NENA Lower Tiyare)	<i>paθəx</i> ‘he opens’, <i>iða</i> ‘hand’ (Borghero 2005)
-------------------------------	--

In other dialect clusters, however, the interdentals *t̥ and *d̥ have been lost. In most cases they have merged with the stops /t/ and /d/. This shift is attested across the Christian NENA dialects of the region from Bohtan and Cudi in the western periphery to Urmī in the eastern periphery, e. g.

Hassana (C NENA Bohtan)	<i>patəx</i> ‘he opens’ (< *pātəx), <i>beta</i> ‘house’ (< *baytā), <i>'ida</i> ‘hand’ (< *'idā)
Bēspən (C NENA Cudi)	<i>patəx, bayta, 'ida</i>
Urmī (C NENA Urmī)	<i>patəx, beta, 'ida</i>

In dialects between these two peripheries, the unvoiced interdental $*\underline{t}$ has undergone other shifts. In the Upper Tiyare dialects $*\underline{t}$ shifts to /š/ in certain contexts, most of which have in common the existence, either synchronically or diachronically, of a preceding high front vowel:

Upper Tiyare	<i>paθəx</i> ‘he opens’ (< * <i>pātəx</i>), <i>beša</i> ‘house’ (< * <i>bay<u>tā</u></i>), <i>mliša</i> ‘full fs.’ (< * <i>mlū<u>tā</u></i>), <i>'iða</i> ‘hand’ (< * <i>īdā</i>) (Talay 2008a: 66–68)
--------------	---

In many dialects in the C Hakkari cluster the interdental $*\underline{t}$ is weakened to /h/ or Ø. When its reflex is Ø, adjacent vowels merge or are connected by a glide, e. g.

Baz (C NENA Hakkari)	<i>maha</i> ‘village’ (< * <i>mā<u>tā</u></i>), <i>xədyuwwa</i> ‘happiness’ (< * <i>hidyu<u>tā</u></i>) (Mutzafi 2000)
Jilu (C NENA Hakkari)	<i>ma</i> ‘village’ (< * <i>mā<u>tā</u></i>), <i>məlkuwwa</i> ‘kingdom’ (< * <i>malkū<u>tā</u></i>) (Fox 1997: 16)

There is a certain degree of lexical and morphological conditioning of this weakening, in that it is a feature of particular lexical items and morphemes. In some lexical items in these dialects the reflex of $*\underline{t}$ is a stop /t/, e. g.

Jilu (C NENA Hakkari)	<i>patəx</i> ‘he opens’ (< * <i>pātəx</i>)
-----------------------	---

In the Jewish NENA dialects of northwestern Iran the interdentals generally shift to the lateral /l/, as is the case in the Jewish trans-Zab NENA dialects of Iraq (chapter 3.4., §5.1):

J. Urmi	<i>paləx</i> ‘he opens’ (< * <i>pātəx</i>), <i>ela</i> ‘festival’ (< * <i>ēdā</i>)
---------	---

In a few words $*\underline{t}$ in this Jewish dialect cluster has lost its oral articulation and shifted to /h/. Most words in which this shift is found have suprasegmental pharyngealization, e. g.

J. Urmi	⁺ <i>ahra</i> ‘town’ (< * <i>a<u>tā</u></i>), ⁺ <i>nahale</i> ‘ears’ (< * <i>natā<u>tā</u></i>)
---------	--

In a few cases, including a number of common words in the lexicon, the reflex of $*\underline{d}$ is /d/:

J. Urmi	<i>ida</i> ‘hand’ (< * <i>īdā</i>), <i>od</i> ‘he does’ (< * <i>āwə<u>d</u></i>)
---------	---

In a small group of NENA dialects in the Bohtan cluster on the western periphery of the region the velar fricative $*\underline{k}$ shifts to a pharyngeal /ħ/:

Hertevin (C NENA Bohtan) *bahe* ‘he weeps’ (< **bākē*)

The dorsal stops **k* and **g* of the *bgdkpt* series undergo palatalization in many of the NENA dialects of the region, in some cases resulting in the affricates /č/ and /j/. In most dialects where palatalization occurs, it is not systematic and tends to be restricted to particular lexical items and phonetic contexts, e. g. Hertevin (C NENA Bohtan) before long vowels and short front vowels (Jastrow 1988: 4). The process is most advanced in the varieties of the C. Urmī dialect that were spoken on the southern Urmī plain, in which affrication of **k* and **g* is regular in all contexts. A similar palatalization of dorsals is found in the Kurdish and Azeri dialects of this area (Khan 2016, vol. 1: 109–112), e. g.

Gulpašan (C NENA Urmī) *malča* ‘king’ (< **malkā*),
jəšra ‘bridge’ (< **gəšra*)

In C. Urmī a dorsal **k* may palatalize completely to /y/ in a few cases, e. g.

**dukta* > *duyta* ‘place’

5.2. Laryngeal Settings of Unvoiced Stops and Affricates

NENA dialects in the region exhibit both aspirated and unaspirated unvoiced stops and affricates. The aspirated series are found elsewhere in NENA but the unaspirated series is unique to this region. The following illustrates the two series in the C. Urmī dialect (the voiced unaspirated consonants are distinguished by a diacritic):

unvoiced aspirated	unvoiced unaspirated	voiced
/p/	/p̪/	/b/
/t/	/t̪/	/d/
/č/	/č̪/	/j/
/c/	/c̪/	/ɟ/

5.3. Pharyngeals and laryngeals

The pharyngeal consonants **ħ* (unvoiced pharyngeal fricative) and **ʕ* (voiced pharyngeal fricative) have been preserved in CNA:

Turoyo *hamro* ‘wine’ (< **hamrā*),
afro ‘dust’ (< **‘aprā*)

In most NENA dialects of the region **ħ* has shifted to /x/, e. g.

C. Urmī *xmara* ‘donkey’ (< **hmārā*)

In a small group of NENA dialects in the Bohtan cluster on the western periphery of the region, the reflex of **h* is /h/ (Jastrow 1994b; Talay 2008a: 44–45). In such dialects a velar fricative **k* has shifted to the pharyngeal /ħ/, so the pharyngeal reflex of **h* may have resulted in the development **h* > **x* > **h*:

Hertevin (C NENA Bohtan)	<i>hmara</i> ‘ass’ (< * <i>hmārā</i>)
--------------------------	--

In all NENA dialects the voiced pharyngeal *‘ has been weakened to the laryngeal // (occasionally /h/) or to Ø

Rumta (C NENA Tiyare)	<i>mar'a</i> ‘illness’ (< * <i>mar'ā</i>), <i>be'e</i> ‘eggs’ (< * <i>bē'ē</i>) (Talay 2008a: 72–78; 2008b)
Qočanəş (C NENA Hakkari)	<i>bihe</i> ‘eggs’ (< * <i>bē'ē</i>) (Talay 2008a: 79)
C. Urmi	+ <i>marra</i> ‘illness’ (< * <i>mar'ā</i>), <i>biyyə</i> ‘eggs’ (< * <i>bē'ē</i>) (Khan 2016, vol. 1: 169–172)

In the variety of the C. Urmi dialect spoken in Armenia a historical *h* shifts to the velar fricative /x/ under the influence of a similar shift in the Armenian dialects of the region, e. g.

C. Urmi (Armenia)	<i>xada</i> ‘thus’ (< <i>hada</i>) (Khan 2016, vol. 1: 104)
-------------------	--

5.4. Pharyngealized consonants

In CNA and the dialects in the western half of the NENA area of the region the historical pharyngealized consonants of earlier Aramaic **t* and **s* have generally been preserved as pharyngealized segments:

Turoyo (CNA)	<i>turo</i> ‘mountain’ (< * <i>tūrā</i>), <i>sawmo</i> ‘fast’ (< * <i>sawmā</i>)
Ashitha (C NENA Tiyare):	<i>tura</i> , <i>soma</i>

In addition, several of these dialects exhibit pharyngealization of a number of other consonants, especially sonorants and labials. These were not historically pharyngealized segments, but now have in some cases developed into fully phonemic consonants that contrast with the non-pharyngealized counterpart. This applies in particular to the rhotic **r*:

Ashitha (C NENA Tiyare)	' <i>amra</i> ‘wool’ (< *' <i>amrā</i>), ' <i>amra</i> ‘she says’ (< *' <i>āmrā</i>)
-------------------------	---

In some of the Tiyare dialects, in addition to the non-pharyngealized alveolar /r/, a non-pharyngealized retroflex /ɻ/ has developed, which both can stand in opposition to the pharyngealized rhotic /ṛ/:

Ashitha (C NENA Tiyare)	<i>p̥aŋa</i> ‘layer of dust’ (< * <i>p̥ārā</i>), <i>para</i> ‘ornamental coin’ (< Turk. <i>para</i>), <i>para</i> ‘lamb’ (< * <i>parrā</i>) (Mutzafi 2014)
-------------------------	---

In the NENA dialects in the eastern half of the region, the original pharyngealization of the consonantal segments has developed into a suprasegmental feature that takes a whole word as its domain (marked here by the symbol ⁺). In words with suprasegmental pharyngealization in Christian dialects, the original pharyngealized consonant segment **t* has become the tense, unaspirated stop segment /t/, which is distinct from the aspirated stop /t/ (Khan 2013; 2016, vol. 1: 107–139), e. g.

C. Urmī	⁺ <i>tuyra</i> ‘mountain’ (< * <i>tūrā</i>) ⁺ <i>soma</i> ‘fast’ (< * <i>ṣawmā</i>)
---------	--

In the Jewish dialects of northwestern Iran, the historical **t* merges with the aspirated stop /t/:

J. Urmī	⁺ <i>tura</i> , ⁺ <i>soma</i>
---------	---

There are differences in the distribution of suprasegmental pharyngealization in the lexicon across the NENA dialects that have this feature (Khan 2016, vol. 1: 75–76), e. g.

C. Urmī	⁺ Mawana (C NENA Šamməsdin-Gawar)
<i>mata</i>	⁺ <i>mata</i>
⁺ <i>dana</i>	<i>dana</i> ‘village’ ‘time’

5.5. Vowels

5.5.1. */ā/ > /o/

The CNA dialects are distinguished from the NENA dialects of the region by the shift of an original long *ā to /o/:

Turoyo	<i>hmoro</i> ‘ass’ (< * <i>hmārā</i>), <i>malko</i> ‘king’ (< * <i>malkā</i>)
--------	--

In most of the NENA dialects of the region, as in NENA dialects in other regions, an original long *ā has not undergone this quality shift, e. g.

Ashitha (C NENA Tiyare)	<i>xmara</i> , <i>malka</i>
C. Urmī	<i>xmara</i> , <i>malca</i>

In a few dialects in the C. Bohtan cluster in the north-west periphery of NENA, which is geographically close to the CNA area, *ā in an open syllable that is not word-final shifts to /o/.

Ruma (C NENA Bohtan)	<i>xmóra</i> ‘donkey’ (< * <i>hmārā</i>), <i>pótax</i> ‘he opens’ (< * <i>pātx̥</i>), but <i>málka</i> ‘king’ (< * <i>malkā</i>), <i>pátxa</i> ‘she opens’ (< * <i>pātxā</i>) (Fox 2009)
----------------------	--

These examples show that the quality shift took place after the shortening of an original long *ā in a closed syllable or an unstressed word-final open syllable, which do not change quality. The quality shift is, therefore, a later phenomenon than in CNA.

In other dialects of the Bohtan cluster, such as Hertevin, a long /a/ is realized as a back [a:] (Jastrow 1988: 15).

5.5.2. Vowel raising

In many NENA dialects of the C Hakkari cluster, a short *a vowel is raised to a higher quality in closed syllables, or syllables that were historically closed:

Jilu (C NENA Hakkari)	<i>kelba</i> ‘dog’ (< * <i>kalba</i>), <i>genawa</i> ‘thief’ (< * <i>gannāwā</i> ‘thief’), <i>keka</i> ‘tooth’ (< * <i>kakkā</i>), <i>breta</i> ‘daughter’ (< * <i>brattā</i>), <i>besima</i> ‘pleasant’ (< * <i>bassima</i>), <i>šedər</i> ‘he sends’ (< * <i>mšaddər</i>) (Fox 1997, 2015)
Dīz (C NENA Hakkari)	<i>kəlba</i> , <i>ginawa</i> , <i>kika</i> , <i>brita</i> , <i>bisima</i> , <i>šidər</i> (Talay 2008a: 140, 240)

The lexical distribution of this vowel-raising differs across the Hakkari dialects, and is found also in isolated cases in adjacent clusters, e. g.

Ashitha (C NENA Tiyare)	<i>genawa</i> , but <i>čaka</i> , <i>čalba</i> , <i>brata</i> , <i>basima</i> , <i>mšadər</i>
C. Urmi	<i>jinava</i> , <i>kika</i> , but <i>calba</i> , <i>brata</i> , <i>basima</i> , <i>+šadər</i>

In the Jilu dialect there is a productive rule whereby a stressed /a/ shifts to /e/ before a following /e/ through a harmony process (Fox 1997: 18–19):

<i>náša</i> ‘man’	<i>néše</i> ‘men’
-------------------	-------------------

In Christian NENA dialects in the eastern area of the region (Hakkari, Šamməs-din-Gawar, Urmi) an original */ē/ is raised to /i/, e. g.

C. Urmi	<i>cipa</i> ‘stone’ (< * <i>kēpā</i>)
---------	--

Concomitantly in these dialects the original back rounded vowel */ð/ is raised to /u/, e. g.

C. Urmi	⁺ <i>natura</i> ‘guard’ (< *nāṭōrā)
---------	--

In these dialects an original */ū/ is fronted to the region of [y], represented here by /ü/:

⁺ Mawana (C NENA Šamməsdin-Gawar)	<i>xabūša</i> ‘apple’ (< <i>habbūšā</i>)
---	---

This seems to have been induced by a similar fronting of *u* in Kurdish dialects of southeastern Turkey and northern Iraq (Haig, this volume, chapter 3.3, §4.1.2.). The shift of *o* > *u* in the NENA dialects can be regarded as a pull-chain effect to fill the space left by the fronting of *u*, which raises also the corresponding front vowel *e* > *i*. In C. Urmi long */ū/ is not fully fronted but shifts to a diphthong with a palatal offglide:

C. Urmi	<i>xabuyša</i>
---------	----------------

In C. Salamas and C. Gawilan, spoken to the north of Urmi, the */ū/ is pushed upwards resulting in the fortition of the offglide of the vowel as a velar fricative:

C. Salamas	<i>xibuxša</i>
------------	----------------

5.6. Diphthongs

The CNA dialects, Turoyo and Mlaḥso, preserve the diphthongs *aw and *ay:

Turoyo/Malḥso	<i>gawzo</i> ‘nut’, <i>qayse</i> ‘wood (pl.)’
---------------	---

Also in some of the NENA dialects of the region the diphthongs are preserved, e. g.

Baz, Maha Xtaya (C NENA Hakkari):	<i>yawna</i> ‘dove’, <i>layša</i> ‘dough’ (Mutzafi 2000: 298)
Halmun (C NENA Tiyare):	<i>mawθa</i> ‘death’, <i>bayθa</i> ‘house’ (Talay 2008a: 145–148)

In many NENA dialects, however, the diphthongs are contracted. The result of the contraction of the diphthong *aw is /o/, e. g.

Ashitha (C NENA Tiyare)	<i>moθa</i> ‘death’ (< * <i>mawtā</i>)
-------------------------	---

The diphthong *ay contracts to various vowels across the NENA area, including /e/, /ɛ/ and /a/. These have a geographical correlation and are characteristic of specific clusters, the vowel /e/ being found in clusters on the eastern side of the region, /ɛ/ in the Tiyare cluster in the central area, and /a/ in the western periphery in the Bohtan and Cudi clusters:

C. Urmi	<i>beta</i> ‘house’ (< *baytā)
Ashitha (C NENA Tiyare)	<i>bəθa</i>
Harbole (C NENA Cudi)	<i>bata</i>

5.7. Syllable structure

The general principles of the syllable structure of the NENA dialects of the region are similar to those of the NENA dialects of Iraq (see chapter 3.4, §5.5). Vowels in open syllables are generally long and those in closed syllables are short. Diacritics are used in the transcription, therefore, only where the length of a vowel deviates from this principle. The /ə/ vowel, however, is always short. A couple of distinctive features of the dialects of the NENA region will be mentioned here.

In some of the NENA dialects of the region, the process of eliminating short vowels in open syllables is more advanced than elsewhere. This involves lengthening short vowels or geminating the following consonants. Such processes are found in cases where short vowels in open syllables arise due to the elision of consonants, e. g.

Baz, Aruntus (C NENA Hakkari)	<i>tára</i> ‘door’ (< *tāra < *tar ‘ā)
C. Urmi	⁺ <i>tárra</i>

It is also found where a short vowel in an open syllable originates as an epenthetic, e. g.

Harbole (C NENA Cudi)	<i>maplīta</i> ‘she brings out’ (< *maplāṭa)
C. Urmi	⁺ <i>maplāṭta</i>

Some dialects add affixed augments to increase the number of syllables in a word. One purpose of this is to increase monosyllabic words to bisyllabic:

Hertevin (C NENA Bohtan)	<i>'et tek</i> ‘there is’ (‘et + augment ek)
Baz, Aruntus (C NENA Hakkari)	<i>'i ni</i> ‘there is’ ('i [< *'ū] + augment ni)

The CNA dialects exhibit a number of differences in syllable structure from NENA dialects.

In Turoyo long vowels are shortened when the syllable is closed, but in Mlahso the vowel generally remains long, or half long (Jastrow 1994a: 23):

Turoyo	Mlahso
<i>dóməx</i> ['do:məx]	<i>doméx</i> [do·'me:x] ‘he sleeps’
<i>dámxi</i> ['dəmxi]	<i>domxi</i> [do·m'xi:] ‘they sleep’

In CNA gemination of consonants is weakened after all vowels, unlike NENA where it is preserved after high front vowels (Khan, this volume, chapter 3.4, §5.5):

Turoyo	Ashitha (C NENA Tiyare)
šato	šata ‘year’
’ezo	’azza ‘goat’

Monosyllabic nouns with an initial consonantal cluster are made bisyllabic by phonemicizing a prosthetic vowel before the cluster, unlike NENA, in which a epenthetic vowel between the consonants of the cluster is phonemicized:

Turoyo	Ashitha (C NENA Tiyare)
’ádmo	dámma ‘blood’ < *dmā
’ášmo	’azza ‘name’ < *šmā

5.8. Stress

In CNA there is a difference in basic stress position between Turoyo and Mlahso, in that Turoyo has penultimate stress where Mlahso has word-final stress:

Turoyo	Mlahso
málko	malkó ‘king’
dóməx	doméx ‘he sleeps’

The attachment of enclitics does not affect stress position:

Turoyo	hárke ‘here’, hárke=yo ‘he is here’
Mlahso	’eskó ‘woman’, ’eskó=yo ‘it is a woman’

The Christian NENA dialects of the region have basic penultimate stress, as is the case with the Christian NENA dialects of Iraq (chapter 3.4, §5.6). The Jewish dialects of northwestern Iran, however, have word-final basic stress:

C. Urmi	J. Urmi
béta	bélá ‘house’
dóməx	+damáx ‘he sleeps’

In the Christian dialects, an original penultimate stress becomes word-final when syllables merge due to loss of consonants. This is particularly conspicuous in dialects of the Hakkari cluster, in which the fricative **t* is lost between vowels. Such a process can lead to phonemic stress oppositions:

Jilu (C NENA Hakkari)	báxta ‘woman’ (< *báxtā)
	baxtá ‘women’ (< *baxtā́tā)

In most of the NENA dialects of the region, as generally in NENA, attachment of the enclitic copula does not affect the stress position:

C. Urmi

basíma + ilə basímələ ‘it is fine’

In Ruma and the related dialects in the Bohtan cluster, however, stress shifts when a copula is added (Fox 2009: 62):

tówa + ile

towéle ‘he is good’

6. Nominal morphology

6.1. Pronouns and copula

The independent pronouns in CNA have the following form:

	Turoyo (Midyat)	Turoyo (villages)	Mlahso
3ms	<i>huwe</i>	<i>hiye</i>	<i>hiye</i>
3fs	<i>hiya</i>	<i>hiya</i>	<i>hiya</i>
3pl	<i>hənne</i>	<i>hənnək</i>	<i>hiyen</i>
2s	<i>hat</i>	<i>hat</i>	<i>hat</i>
2pl	<i>hatu</i>	<i>hatu</i>	<i>hatun</i>
1s	<i>'úno</i>	<i>'ono</i>	<i>'ono</i>
1pl	<i>'ahna</i>	<i>'ahna</i>	<i>'eləna</i>

These pronouns have undergone a number of innovations in comparison with the forms that should be reconstructed as their historical ancestors (Jastrow 1990). The 3rd person forms, for example, are augmented by additional pronominal suffixes: Turoyo (Midyat) *huwe* < **hu* + *e*, *hiya* < **hi* + *a*, *hənne* < **hen* + *hen*. In Mlahso all the 3rd person pronouns have the same stem (originally that of the 3fs pronoun) and the number and gender is only distinguished by the augment. The form of the 1pl pronoun in Mlahso, *eləna*, is originally an accusative form.

The original form of the 3rd person pronouns are used as proclitics with the function of a definite article:

Turoyo	<i>'ú-malko</i>	‘the king’	< * <i>hū malkā</i>
	<i>'i-malakθo</i>	‘the queen’	< * <i>hī-maləktā</i>
	<i>'ám-malke</i>	‘the kings’	< * <i>hen-malkē</i>

In CNA the copula is expressed by enclitic forms of the pronouns:

Turoyo (Midyat)

	Independent pronoun	Enclitic pronominal copula
3ms.	<i>huwe</i>	= <i>yo</i>
3fs.	<i>hiya</i>	= <i>yo</i>
3pl.	<i>hənne</i>	= <i>ne</i>
2ms.	<i>hat</i>	= <i>hat</i>
2fs.	<i>hat</i>	= <i>hat</i>
2pl.	<i>hatu</i>	= <i>hatu</i>
1s	<i>'ūno</i>	= <i>no</i>
1pl.	<i>'aḥna</i>	= <i>na</i>

In the villages there is gender distinction in the 2s:

Turoyo (Midin)

	Independent pronoun	Enclitic pronominal copula
2ms.	<i>hat</i>	= <i>hət</i>
2fs.	<i>hat</i>	= <i>hat</i>

Turoyo has an alternative copula paradigm in which the enclitic pronouns are attached to the existential particle *kit* ‘there is’:

Turoyo (Midin)

3ms.	<i>kətyo</i>
3fs.	<i>kətyo</i>
3pl.	<i>kətnə</i>
2ms.	<i>kəthət</i>
2fs.	<i>kəthat</i>
2pl.	<i>kəthatu</i>
1s	<i>kətno</i>
1pl.	<i>kətna</i>

Deictic pronouns are suffixed to the nouns they modify. The paradigm in Turoyo is as follows:

Near deixis

ms	<i>-ano</i>	‘this’	cf. Syriac <i>hānā</i>
fs.	<i>-aθe</i>	‘that’	cf. Syriac <i>hāde</i>
pl.	<i>-ani</i>	‘these’	cf. Syriac <i>hālēn</i>

Far deixis

ms	<i>-awo ~ -awok</i>	‘that’	cf. Syriac <i>haw</i>
fs.	<i>-ayo ~ -ayək</i>	‘that’	cf. Syriac <i>hāy</i>
pl.	<i>-anək</i>	‘those’	cf. Syriac <i>hānōn</i> (m.) <i>hānēn</i> (f.)

In Turoyo the suffixed deictics are combined with the prefixed definite article, e. g. '*u-malk-anō* ‘this king’, '*u-malk-awō* ‘that king’ (Jastrow 1985: 39). In Mlahso the definite article is not obligatory in such constructions (Jastrow 1994a: 32).

Turoyo also has a deictic copula which points to the existence of an entity in the speech situation, formed from the particle *ka* + pronominal object suffixes containing the element *l*. The pronominal copula may be optionally added to some forms (Jastrow 1985: 122):

Turoyo (Midin)

3ms.	<i>kalé</i> ~ <i>kalé=yo</i>	‘There he is’
3fs.	<i>kalá</i> ~ <i>kalá=yo</i>	‘There she is’
3pl.	<i>kalán</i> ~ <i>kalán=ne</i>	‘There they are’, etc.
2ms.	<i>kaláx</i>	
2fs.	<i>kaláx</i>	
2pl.	<i>kalóxu</i>	
1s.	<i>kali</i> ~ <i>kali=no</i>	
1pl.	<i>kalán</i>	

The NENA dialects of the region exhibit considerable diversity in the forms of the personal pronouns and deictic pronouns. There are many parallels with the forms of the pronouns used in the NENA dialects in Iraq (chapter 3.4, §6.1). Here we shall focus on the differences between NENA and CNA.

In all NENA dialects the 3rd person singular personal pronouns, i. e. the pronouns that have an anaphoric function, have a prefixed *'a-*. The proto-NENA forms can be reconstructed as *a-hu* (3ms) and *a-hi* (3fs) (Hoberman 1988). These forms have been preserved in some dialects of the region:

Turoyo (Midyat)		Baz, Maha Xtaya (C NENA Hakkari)
3ms.	<i>huwe</i>	<i>'ahu</i>
3fs.	<i>hiya</i>	<i>'ahi</i>

The forms *'ahu* and *'ahi* may ultimately go back to far deictic pronouns **hāhū* and **hāhī*, in which the **hā-* is a deictic element. These deictic pronouns would subsequently have taken on the function of anaphoric pronouns. Alternatively, the *'a-* prefix may have developed by analogy with 1st and 2nd person pronouns, which also begin with *'a-*.

The reconstructed proto-NENA singular deictic pronouns and the reflexes of deictic pronouns in a selection of NENA dialects are presented in the table below. This includes also the anaphoric pronoun, i. e. the 3rd person personal pronoun:

	Proto-NENA	Hertevin (Bohtan)	Baz Aruntus (Hakkari)	Ashitha (Tiyare)	C. Urmi	J. Urmi
Near deixis						
ms	<i>hādā</i>	'ohá	'aha	'awwa	'aha	ya
fs	<i>hādī</i>	'ehá	'aha	'ayya	'aha	ya
Far deixis						
ms	<i>hāhū</i>	'awa	'awaha	'awaha	+ 'avva	o
fs	<i>hāhī</i>	'aya	'ayaha	'ayaha	'ayya	o
Anaphoric						
ms	<i>hāhū</i>	'ahu	'aw	'o	+ 'av	o
fs	<i>hāhī</i>	'ahu	'ay	'ɛ	'ay	o

As can be seen, various innovations have taken place across the dialects. In the selection of dialects in the table, J. Urmi retains the original configuration whereby the same pronoun expresses far deixis and anaphora (*o* is a contraction of **hāhū*, which in J. Urmi is of common gender). Other dialects innovated by distinguishing the anaphoric pronoun from the far deixis pronoun by the suffixing of the deictic element *-ha* to the latter, e. g. C. Urmi + 'av (anaphoric), + 'avva (far deixis < *'awha). Some dialects preserved a form of the near deixis pronoun that is derived ultimately from the proto-form, e. g. Baz, Maha Xtaya 'aha < **hādā*. Many dialects, however, replaced the original near deixis form by one based on the stem of the anaphoric pronoun, by adding the deictic element *-ha* to this stem, e. g. Ashitha 'awwa (< *'aw-ha). It is likely that such forms originated as far deixis pronouns (cf. C. Urmi + 'avva), but shifted to near deixis forms when a further innovation occurred in the far deixis forms involving the addition of a second deictic *-ha* suffix, as is the case in the Ashitha dialect: 'awaha (< *'aw-ha-ha). Some dialects such as Hertevin express the distinction between the near deixis and far deixis form by differences in the contraction of a diphthong ('ohá ms. near deixis < *'awha, 'awa ms. far deixis < *'awha).

As in the NENA dialects of Iraq, many dialects can express 'very far' deixis by phonological strengthening, e. g.

Hassana (C NENA Cudi)

	Far deixis	Very far deixis
ms	'awáha	'awá'ha
fs.	'ayáha	'ayá'ha

NENA dialects of the region have an enclitic copula, as in NENA dialects in other regions (Khan, this volume, chapter 3.4, §7, chapter 4.4, §7.1). These originated historically as enclitic pronouns, as in Turoyo, but in most dialects they have undergone morphological change resulting in their assimilation to the inflection

of presentative constructions and the inflections of verbs (Khan 2001, 2012). The most archaic NENA dialects in this respect are on the western periphery, in which the 1st and 2nd person copula forms are still recognizable as enclitic forms of the independent pronouns, e. g.

Hertevin (C NENA Bohtan)

	Independent Pronoun	Enclitic Copula
3ms	'ahu	=ile
3fs.	'ahu	=ila
3pl.	'ahni	=ini
2ms.	'ahət	=ihət
2fs.	'ahat	=ihat
2pl.	'ahniton	=əhton
1ms.	'ana	=ina
1fs.	'ana	=ina
1pl.	'ahnah	=əhnah

As with the NENA dialects of Iraq, most NENA dialects in the region have a deictic copula (chapter 3.4, §7) in addition to an enclitic copula. A feature of some NENA dialects of the region is the existence of a near deixis and a far deixis series of copulas, e. g. in C. Urmi, where the far deixis series is restricted to the 3rd person:

	Near deixis	Far deixis	Enclitic
3ms.	dulə	velə	=ilə
3fs.	dula	vela	=ila
3pl.	duna	vena	=ina
2ms.	duvət		=ivət
2fs.	duvat		=ivat
2pl.	dutun		=itun
1ms.	duvən		=ivən
1fs.	duvan		=ivan
1pl.	duvax		=ivəx

6.2. Pronominal possessive suffixes

The pronominal suffixes of the NENA dialects of the region exhibit a similar range of forms as is found in the NENA dialects of Iraq (chapter 3.4, §6.1). Here we shall concentrate on the pronominal possessive suffixes in the CNA dialects and their relationship to the NENA suffixes.

Turoyo has two series of suffixes. Series I consists of the basic suffixes and series 2 consists of suffixes that are composed of the morpheme *-ayð-* followed by a suffix

	Series 1	Series 2
3ms.	<i>-e</i>	<i>-ayðe</i>
3fs.	<i>-a</i>	<i>-ayða</i>
3pl.	<i>-ayye</i>	<i>-aθθe</i>
2ms.	<i>-əx</i>	<i>-ayðəx</i>
2fs.	<i>-ax</i>	<i>-ayðax</i>
2pl.	<i>-ayxu</i>	<i>-aθxu</i>
1fs.	<i>-i</i>	<i>-ayði</i>
1pl.	<i>-an</i>	<i>-ayðan</i>

The first series of suffixes without the augmented morpheme is used predominantly to express inalienable possession, e. g. with kinship terms and parts of the body. The second series, with the heavier suffix, is used as a general possessive suffix. The second series is combined with the definite article, whereas the first series combines without the article:

- 'em-e 'his mother'
 'u-bayt-ayðe 'his house'

The heavier series derives from the contraction of an originally independent pronominal genitive particle with a noun < *'u-bayto-diðe. A similar distinction between a light and a heavy series of possessive suffixes is found in Mlahso, in which the inserted element in the heavy series has the form *-ez-* (Jastrow 1994a: 29).

In NENA dialects a genitive independent pronoun can optionally be used, typically to express alienable possession, e. g. C. Urmi *betət diyyan* 'our house'. Whereas in NENA this construction has the status of a pragmatic strategy, in CNA it has become grammaticalized and the genitive pronoun has become morphologically bonded with the head noun.

6.3. Inflection of nouns and adjectives

The inflection of nouns and adjectives in the NENA dialects of the region follow the same basic patterns as the NENA dialects of Iraq (chapter 3.4, §6.3). Due to the loss of intervocalic **t̪* in the Christian dialects of the Hakkari cluster, however, nominal inflections that originally contained this consonant are contracted, e. g.

Baz, Maha Xtaya (C NENA Hakkari) (Mutzafi 2000: 303)	šónā 'sleep' < *šántā
	baxtá 'women' < *baxtátā
	xawá 'sisters' < *hatwáṭā

In the eastern sector of the region, the common nominal plural ending **-ē* has been analogically extended to the feminine plural ending **-ātā*, resulting in the form *-ate* or *-atə*, e. g.

⁺ Mawana (Šamməsdin-Gawar)	C. Urmi	Ashitha (C Tiyare)
<i>naš-e</i>	<i>naš-ə</i>	<i>naš-e</i>
<i>baxt-ate</i>	<i>baxt-atə</i>	<i>baxt-aθa</i>

‘people’
‘women’

The CNA dialects exhibit inflections of nouns and adjectives that are comparable to NENA. This includes the analogical extension of the basic plural ending *-e* to the feminine ending in some dialects:

	Singular	Plural
Turoyo (Midyat)	<i>karmo</i> ‘vineyard’ <i>'ar'o</i> ‘field’	<i>karme</i> <i>'ar'oθo</i> (< *'ar'ātā)
Turoyo (villages)	<i>karmo</i> ‘vineyard’ <i>'ar'o</i> ‘field’	<i>karme</i> <i>'ar'oθe</i>
Mlaḥso	<i>karmo</i> ‘vineyard’ <i>'ar'o</i> ‘field’	<i>karme</i> <i>'ar'ose</i>

Some of the Jewish dialects of northwestern Iran spoken to the south of Urmi have borrowed a definite article suffix from Sorani Kurdish with the form *-āke*, e. g. J. Šəno *bela* ‘house’, *belāke* ‘the house’. This is used also in the trans-Zab Jewish dialects of Iraq (chapter 3.4, §6.1) and western Iran (chapter 4.4, §6.3.1).

6.4. Nominal annexation

In Turoyo the annexation of a head noun to a modifying dependent noun is expressed by prefixing the inherited Aramaic subordinating particle *d* to the dependent noun, as in earlier Aramaic (Jastrow 1985: 44–47). If the nouns of the construction are definite, they have the definite article (see §6.1):

<i>'i-barθo</i>	<i>d-u-malko</i>
ART.FS-daughter	GEN-ART.MS-king
‘the daughter of the king’	

Another construction is used in Turoyo whereby a 3ms pronominal suffix is placed on the head noun anticipating the dependent noun:

<i>foθ-e</i>	<i>d-u-babo</i>
face-his	GEN-ART.MS.-father
‘the face of the father’	

This 3ms is invariable and is used also when the dependent noun is feminine singular or plural:

<i>foθ-e</i>	<i>d-i-emo</i>
face-his	GEN-ART.FS.-mother
‘the face of the mother’	

In most NENA dialects of the region the subordinating particle *d* is suffixed to the head noun and devoiced, e. g.

C. Urmī	
<i>bet-ət</i>	<i>malca</i>
house-GEN	king
‘the house of the king’	

Such constructions may derive historically from constructions with a fossilized 3ms anticipatory pronoun, as is found in Turoyo, as already suggested by Nöldeke (1868: 149):

<i>bet-ə-t</i>	<i>malca</i>
house-his-GEN	king

The suffix may be contracted in closely-knit phrases, in particular those expressing kinship relations or inalienable possession:

C. Urmī	
<i>brūn</i>	<i>malca</i>
son	king
‘the son of the king’	

In the Jewish dialects on the eastern periphery an *ay* element is often placed between the two nouns. The etymology of this is likely to be an Aramaic demonstrative pronoun, but its usage was probably induced by the *izafe* particle of Iranian languages in contact (Khan 2008: 176–178):

<i>lel-ət</i>	<i>ay-xlula</i>
night-GEN	DEM-wedding
‘the night of the wedding’	

As in the NENA dialects of Iraq (chapter 3.4, §6.4), the genitive particle *d* has become bonded with the NENA demonstrative pronouns to form genitive allo-morphs of the pronouns (Khan 2016: 239–242):

C. Urmī	
<i>bet-ət</i>	<i>d-o naša</i>
house-GEN	GEN-that man
‘the house of that man’	

7. Verbal morphology

The basic principles of verbal derivational morphology and inflection in the NENA dialects of the region are the same as those that have been described for the NENA dialects of Iraq (chapter 3.4, §7). Here we shall concentrate on some distinctive features of the NENA dialects of the region and on the verbal morphology of the CNA dialects.

The three derivational stems of verbs are preserved in the Christian NENA dialects, e. g.

Ashitha (C NENA Tiyare)

		Present Base	
Stem I	<i>p-θ-x</i>	<i>paθəx</i>	'to open'
Stem II	<i>mb-š-l</i>	<i>mbašəl</i>	'to cook'
Stem III	<i>m-pl-x</i>	<i>mapləx</i>	'to use'

In some Christian NENA dialects, predominantly in the eastern sector of the region, the *m-* prefix of stem II is lost, resulting in a merger of stem II with stem I in the present base:

C. Urmi

Stem I

p-t-x 'to open'

Present base:	<i>patəx</i>
Past base:	<i>ptix-</i>
Resultative participle:	<i>ptixa</i>
Imperative:	<i>ptux</i>
Infinitive:	<i>ptaxa</i>

Stem II

b-š-l 'to cook'

Present base:	<i>bašəl</i>
Past base:	<i>bušəl</i>
Resultative participle:	<i>bušla</i>
Imperative:	<i>bašəl</i>
Infinitive:	<i>bašulə</i>

Stem III

m-pl-x 'to use'

Present base:	<i>mapləx</i>
Past base:	<i>mupləx</i>
Resultative participle:	<i>mupləxxa</i>
Imperative:	<i>mapləx</i>
Infinitive:	<i>mapluxə</i>

In the Jewish dialects of northwestern Iran the merger of the stems and their internal vocalic patterns is more advanced than in the Christian dialects of the region. In J. Urmī, for example, the original stem II has been lost completely and the vocalic patterns of the surviving two stems have been levelled.

Stem I

⁺*q-t-l* ‘to kill’

Present base:	⁺ <i>qatəl-</i>
Past base:	⁺ <i>qtəl-</i>
Resultative participle:	⁺ <i>qtila</i>
Imperative:	⁺ <i>qtul</i>
Infinitive:	⁺ <i>qatole</i>

Stem II

m-nx-p ‘to shame’

Present base:	<i>manxəp</i>
Past base:	<i>mənxəp-</i>
Resultative participle:	<i>mənxipa</i>
Imperative:	<i>mənxup</i>
Infinitive:	<i>manxope</i>

In most of the NENA dialects across the region the resultative perfect of both transitive and intransitive verbs is expressed by combining the resultative participle with the copula (see chapter 3.4, §7). This construction, however, is not used in some of the dialects of the Bohtan cluster in the western periphery of the region. Instead they express the perfect by inflections of the past base, the perfect being distinguished from the perfective by different inflectional suffixes. In Ruma and the related dialects described by Fox (2009) the perfective is expressed by the past base inflected by L-suffixes and the perfect by the past base inflected by D-suffixes (see chapter 3.4, §7 for these inflectional suffixes):

	Perfective	Perfect
Transitive:	<i>grəš-le</i> ‘He pulled’	<i>griš-Ø</i> ‘He has pulled’
	<i>grəš-la</i> ‘She pulled’	<i>griš-a</i> ‘She has pulled’
Intransitive:	<i>qəm-le</i> ‘He stood up’	<i>qim-Ø</i> ‘He has stood up’
	<i>qəm-la</i> ‘She stood up’	<i>qim-a</i> ‘She has stood up’

In Hertevin only the perfect of intransitive verbs can be expressed in this way (Jastrow 1988: 46–59):

	Perfective	Perfect
Transitive:	<i>greš-le</i> ‘He pulled’	—
	<i>greš-la</i> ‘She pulled’	
Intransitive:	<i>qəm-le</i> ‘He stood up’	<i>qəm-Ø</i> ‘He has stood up’
	<i>qəm-la</i> ‘She stood up’	<i>qim-a</i> ‘She has stood up’

In such a system, in which the perfect form is less developed, the function of the transitive perfective form *grāšle* is extended to cover also that of the transitive perfect.

In the Jewish dialects of northwestern Iran on the eastern periphery of the region one finds a system similar to that of Hertevin with a perfect constructed with S-suffixes on the past base of intransitive verbs. The perfect of transitive verbs, however, is expressed by a resultative participle with a copula (Khan 2008: 70–80):

	Perfective	Perfect
Transitive:	<i>grāš-le</i> 'He pulled'	<i>grišé</i> 'He has pulled' (< * <i>griša</i> + <i>ile</i>)
	<i>grāš-la</i> 'She pulled'	<i>grāštá</i> 'She has pulled' (< * <i>grāšta</i> + <i>ila</i>)
Intransitive:	<i>qəm-le</i> 'He stood up'	<i>qim-Ø</i> 'He has stood up'
	<i>qəm-la</i> 'She stood up'	<i>qim-a</i> 'She has stood up'

In many NENA dialects of the region transitive perfective verbs express definite objects by absolute nominative D-suffixes, which occur before the L-suffixes, e. g.

C. Ashitha (Borghero 2005)

<i>griš-a-le</i>	<i>griš-i-le</i>
pull.PFV-D.3FS-L.3MS	pull.PFV-D.3PL-L.3MS
'he pulled her'	'he pulled them'

In such dialects the D-suffix agrees with definite nominal objects in the clause, as in the NENA dialects Iraq (chapter 3.4, §7).

Some dialects do not allow any absolute pronominal objects at all and express all pronominal objects with accusative L-suffixes. This applies, for example, to the Ruma dialect (C NENA Bohtan) (Fox 2009: 53). In some dialects that have this pronominal object-marking strategy, adjustments are made to the form of the L-suffix that marks the subject. A noteworthy adjustment of this kind has been documented in the Hertevin dialect (C NENA Bohtan) by Jastrow (1988: 61). In this dialect the L-suffixes remain unadjusted if the subject is third person but are adjusted by giving them the ending of D-suffixes if the subject is 1st or 2nd person, e. g.

C Hertevin

(a) Perfective stem + L-suffix (A)

<i>hze-le</i>	'he saw'
<i>hze-lox</i>	'you (ms) saw'
<i>hze-li</i>	'I saw'

(b) Perfective stem + L-suffix (A) + L-suffix (O)

<i>hze-le-li</i>	'he saw me'
<i>hze-let-ti</i> (< <i>hze-let-li</i> , cf 2ms D-suffix <i>-et</i>)	'you saw me'
<i>hze-len-ne</i> (< <i>hze-len-le</i> , cf 1ms D-suffix <i>-en</i>)	'I saw him'

The result is the conversion of the 1st and 2nd person L-suffixes that express the agentive subject (A) to suffixes that resemble the form of nominative D-suffixes.⁴

In most NENA dialects of the region the progressive is expressed by constructions consisting of the infinitive combined with the copula, which appears to have developed under the influence of the model of progressive forms based on the infinitive in Eastern Armenian. This construction is not used, however, in some of the dialects of the Bohtan cluster on the western periphery, e. g. Ruma and Hertevin. In these dialects inflections of the present base are used to express the progressive (Fox 2009: 31, 55; Jastrow 1988: 54).

As in the NENA dialects of northern Iraq, in the NENA dialects of the region preverbal particles are attached to the present base when it expresses indicative realis mood. In most dialects of south-eastern Turkey this is *'i-* or *y-*. In some dialects the prefix is elided in verbs beginning with a strong consonant but is generally preserved before verbs beginning with a vowel, e. g.

C. Ruma (Bohtan) (Fox 2009: 47)

y-oxəl ‘he eats’ (realis)

oxəl ‘may he eat’ (irrealis)

In dialects in which the progressive is expressed by a construction based on the infinitive, present forms with such a prefix express the habitual. In dialects that do not use such a progressive construction, the form with the prefix may express both the habitual and the progressive. In the NENA dialects of southeastern Turkey there are traces of a *k*- prefix, originally, it seems, combined with the prefix *'i*, e. g. Mazra (Txuma) *k-i-'*odax ‘we do’. In the Upper Tiyari dialects this has become restricted to a few lexical verbs and palatalized to *č*-, e. g. *či* ‘in’ (< **k-i-yād* ‘on’) (Talay 2008b: 306). In the NENA dialects of northwestern Iran, the *k*-element has been preserved to a greater extent. In C. Urmī the indicative prefix has the form *ci-* (< *k + i*), e. g. *ci-patəx* ‘he opens’. In C. Urmī an original velar **k* has been fronted to the palatal /c/. J. Urmī has preserved *k*- before verbs beginning with weak consonants, e. g. *k-xəl* ‘he eats’ (< *k + axəl*).

The future is expressed throughout the NENA dialects of the region by a prefix particle deriving from the volitive expression **bā'ē d-* ‘wants to’. This has the form *bəd-*, which is often contracted to *b-* or, before vowels, *p̪t-* /t/.⁵ In J. Urmī the future may be expressed also by *gbe*, the indicative 3ms form of ‘to want’, e. g. *gbe garšen* ‘I shall pull’ (Khan 2008: 76).

Some dialects of the region use the particle *de-* before an imperative to give immediacy and intensity to the command, e. g. Ashitha (C. Tyare) *de-plut!* ‘go out!’ (Borghero 2005: 126).

⁴ For further details of the various constructions with transitive verbs in the NENA dialects of the region see Khan (2017).

⁵ The sound transcribed as /t/ is an unaspirated unvoiced alveolar stop. A /t/ without a diacritic represents an aspirated unvoiced alveolar stop.

The CNA dialects are similar to the NENA dialects in that their verbal system is largely constructed on bases that were active and passive participles of earlier Aramaic, inflected with suffixes corresponding to the D-suffixes and L-suffixes of NENA. The CNA dialects, however, exhibit a number features that are distinctive from NENA and these will be the main focus of the following discussion.

As in most NENA dialects Turoyo has three derivational stems, e. g.

Stem I	<i>n-š-q</i> ‘to kiss’	
	Present base:	<i>nóšaq-</i>
Stem II	<i>mḥ-l-q</i> ‘to throw’	
	Present base:	<i>mḥáləq-</i>
Stem III	<i>m-nš-f</i> ‘to cause to dry’	
	Present base:	<i>mánšəf</i>

Unlike NENA, however, Turoyo also has a passive base for each stem:

Stem I	Passive <i>f-t-h</i> ‘to open’	
	Present base:	<i>máftəh-</i>
Stem II	Passive <i>mḥ-l-q</i> ‘to throw’	
	Present base:	<i>mihátləq-</i>
Stem III	Passive <i>m-nš-f</i> ‘to cause to dry’	
	Present base:	<i>mitánšəf-</i>

These passive bases are derived historically from passive and middle derivational patterns that contained a *t* prefix in earlier Aramaic (cf. the so-called *'etpe 'el*, *'etpa* ‘al and *'ettaþ* ‘al forms in Syriac). The *t* is preserved only in stem III.

The past bases are inflected with L-suffixes and D-suffixes. The forms of these in Tuoryo are as follows (Jastrow 1985, 128–129):

L-suffixes		D-suffixes
3ms.	<i>-le</i>	<i>-Ø</i>
3fs.	<i>-la</i>	<i>-o</i>
3pl.	<i>-:e</i> (after consonant), ⁶ <i>-lle</i> (after vowel)	<i>-i</i>
2ms.	<i>-ləx</i>	<i>-ət</i>
2fs.	<i>-lax</i>	<i>-at</i>
2pl.	<i>-xu</i> (after consonant) <i>-lxu</i> (after vowel)	<i>-utu</i>
1ms	<i>-li</i>	<i>-no</i>
1fs.	<i>-li</i>	<i>-ono</i>
1pl.	<i>-lan</i>	<i>-ina</i>

⁶ I. e. the final consonant of the verbal base is geminated.

The L-suffixes are used with transitive verbs and the D-suffixes with intransitive and passive verbs. The distribution of the suffixes, therefore, is based on transitivity and voice. Intransitive verbs nearly all belong to stem I, where they have a past base that is distinct in form from the past base of active stem I verbs. The two bases are derived historically from different participle forms of earlier Aramaic (transitive < the passive participle **CCīC* and intransitive < the verbal adjective **CaCCīC*). The category of intransitive includes unaccusative verbs and also verbs with non-agentive experiencer subjects (e. g. ‘to hear’, ‘to know’) (Furman and Loesev 2015):

Stem I Active transitive

Past base:	<i>nšáq-le</i>	‘He kissed’
	<i>nšáq-la</i>	‘She kissed’

Stem I Intransitive

Past base:	<i>dámāx-Ø</i>	‘He slept’
	<i>damíx-o</i>	‘She slept’

The dialect of Mlahso has a system of derivational stems that is similar to that of Turoyo but Mlahso exhibits significant differences from Turoyo in the distribution of the inflectional suffixes that are added to past bases. In Mlahso the distribution of the suffixes correlates with verbal aspect rather than transitivity and voice. The D-suffixes are used on past bases to express the resultative perfect of both intransitive and transitive verbs. This construction is attested only in stem I on past bases equivalent to the intransitive past base of Turoyo (Jastrow 1994a: 33, 66):

n-f-q (stem I) ‘to go out’

<i>nafiq-Ø</i>	‘He has gone out’
<i>nafiq-o</i>	‘She has gone out’

s-y-m (stem I) ‘to do’

<i>saym-Ø</i>	‘He has made’
<i>saym-o</i>	‘She has made’

L-suffixes are used on transitive, intransitive and passive verbs to express a perfective aspect. This configuration of D-suffixes and L-suffixes has parallels with the verbal system of some NENA dialects in the Bohtan cluster. Neither Turoyo nor Mlahso have constructions consisting of the resultative participle + copula or infinitive + copula to express the resultative perfect and progressive respectively. This differs from most of the NENA dialects of the region, but, again, coincides with the verbal system of dialects in the NENA Bohtan cluster.

Turoyo expresses 3rd person pronominal objects of transitive verbs by absolute D-suffixes as in many NENA dialects, e. g.

grəš-Ø-li
pull.PST-D.3MS-L.1S
'I pulled him'

griš-o-li
pull.PST-D.3FS-L.1S
'I pulled her'

griš-i-li
pull.PST-D.3PL-L.1S
'I pulled them'

The profile of Turoyo with regard to pronominal object-marking is, therefore, similar to that of many NENA dialects. Mlahso, by contrast, does not allow any absolute pronominal objects and all pronominal objects are expressed by accusative L-suffixes or separate accusative prepositional phrases (Jastrow 1994a: 54–56).

A point of difference between Turoyo and NENA is that in transitive clauses verbs do not have pronominal suffixes agreeing with a definite object argument (Hemmauer and Waltisberg 2006: 32, 35) and as a result there is no absolute agreement in verbs derived from past bases:

g-nəšq-o *'u-zlām*
FUT-kiss.PRS-D.3FS ART.MS-man
'She will kiss the man'

nšəq-le *'i-aθto*
kiss.PST-L.3MS ART.FS-woman
'He kissed the woman'

Turoyo expresses the progressive and perfect by combining prefixes with the present and past base respectively, e.g. *kal-ko-nošəq* 'he is kissing', *ko-qayəm-Ø* 'he has arisen' (Jastrow 1985: 146–150). The future is expressed in Turoyo by attaching to the present base a preverbal particle with various allomorphs (*g*, *k*, *gd*, *kt*), all of which derived historically from a particle with the form **gəd* (Jastrow 1985: 148–149). This can also express habitual aspect. In Mlahso the future particle has the form *d-*, which is likely to be related (Jastrow 1994a: 52). The particle **gəd* may be compared to the particle *ket-* which is used in the Hertevin (C NENA Bohtan) with the sense of 'to be able'.

A distinctive feature of Turoyo that is not found in the NENA dialects of the region is the construction used to express the two pronominal objects of ditransitive verbs. In these constructions the pronominal indirect object is expressed by an L-suffix and the pronominal direct object is expressed by an enclitic pronoun that has the same form as the copula. This expression of the pronominal direct object is restricted to the 3rd person:

Turoyo (Jastrow 1985: 142)

gd-oba-l-lax=yo

FUT-give.PRS-D.1FS-L.2MS=3S

‘I shall give him/her/it to you’

Parallel constructions are found in some of the Christian NENA dialects of the Mosul plain (chapter 3.4, §7).

8. Syntax of basic clauses

8.1. Copula clauses

The syntax of copula clauses in NENA has been described in chapter 3.4., §8.1. Here we shall only make the observation that, as in the NENA dialects of Iraq, the moveability of the copula from its default position on the predicate differs across the dialects. The general observation to be made is that in the Christian NENA dialects of the region the copula may be moved and placed after constituents other than the predicate in order to express the prominence of the constituent. In the Jewish dialects of northwestern Iran, however, the copula is fixed on the predicate, as is the case with the trans-Zab Jewish dialects of Iraq:

- (1) J. Urmī (Khan 2008: 316)⁷

*dad-óx +tajàr=ile?| amon-i +tajér=ile| lá
father-your merchant=COP.3MS uncle-my merchant=COP.3MS NEG
dad-i.|
uncle-my*

‘Is your father a merchant?’ ‘*My uncle* is a merchant, not my father.’

Contrast C. Urmī, where such movement is possible (Khan 2016, vol. 2: 296):

- (2) *Yósəp +xòr-an=ila.| là,| Yònán=ilə +xór-an.|*

*Yosəp friend-our=COP.3ms NEG Yonan=COP.3MS friend-our
‘Yosəp is our friend.’ ‘No, Yònán is our friend.’*

The movement of the copula from the predicate position to mark the prominence of other constituents in a clause is possible also in Turoyo (3a–b) on the western periphery of the region:

⁷ In these examples intonation group boundaries are marked by |, the nuclear stress of the intonation group is marked by a grave accent and non-nuclear stress is marked by acute accents.

- (3) a. *ahún-ux* *hárke=yo.*¹
 brother-your.MS here=COP.3S
 My brother is here.
 b. *lò,*¹ *ḥòθ-i=yo* *hárke,*¹ *ló* *ahún-i.*¹
 No sister-my=COP.3S here not brother-my
 ‘No, my sister is here, not my brother.’

For other contexts where the copula is moved in Turoyo see Waltisberg (2016: 97–99).

In the NENA dialects the copula is negated by prefixing to it the negative particle *la*, e. g. C. Ashitha *lele* ‘he is not’ (<*la* + *ile*). In the Jewish dialects of northwestern Iran, the stem of the copula is different when negated, e. g. J. Urmia *ile* ‘he/it is’, *lewe* ‘he/it is not’. In Turoyo the negator of verbs is *lo*, but this is not used with the enclitic pronominal copula. Rather the negative copula is expressed by an alternative paradigm consisting of the negative existential particle *lat* and the enclitic copula, e. g. *-yo* ‘he/she/it is’, *latyo* ‘he/she/it is not’.

8.2. Verbal clauses

The description of the ordering of constituents in the verbal clauses of the NENA dialects of Iraq (chapter 3.4, §8.2) apply also to the NENA dialects of the region with which this chapter is concerned. As in Iraq, the main typological split is between the Jewish trans-Zab dialects in northwestern Iran and the Christian dialects. The Jewish trans-Zab dialects of the region exhibit the same type of word order as those of Iraq, the most distinctive feature of which is the regular placement of the direct object before the verb. The Christian NENA dialects of the region exhibit a more flexible word order with the placement of the direct object after the verb being the unmarked order, as is the case with the Christian NENA dialects of Iraq. Flexible order with the unmarked positioning of the object after the verb is also the typology of the ordering of constituents in Turoyo (Waltisberg 2016: 249–254).

In Turoyo verbs are negated by the particle *lo*, which contracts to *la* in closed syllables. It is placed before all preverbal particles, including the future/habitual particle (Jastrow 1992: 29):

- (4) *lo ko-do 'ər-Ø*
 NEG IND-return.PRS-D.3MS
 ‘He is not returning’
- (5) *la-g-šomə '-Ø*
 NEG-FUT-hear.PRS- D.3MS
 ‘He will not her’/‘He does not hear (habitual)’

In Turoyo the negator cannot be used with the imperative. A prohibition is expressed by negating the irrealis present, e. g. *ḥər!* ‘Look’, *lò-ḥayrət!* ‘Do not look (ms.)!’

In the NENA dialects of the region verbs are negated by the particle *la*, e. g.

- (6) C. Ashitha (Tiyare)
la-šate-Ø
 NEG-drink.PRS-D.3MS
 ‘He does not drink’.

In most NENA dialects the future particle cannot be used with the negator, so a negated future is identical to the present habitual, e. g.

- (7) C. Ashitha (Tiyare) (Borghero 2005: 129)
la-šate-Ø
 NEG-drink.PRS-D.3MS
 ‘He will not drink’.

In dialects that express the habitual by a preverbal particle *i-/y*, the sequence *la* + *y* contracts to *le*, e. g.

- (8) C. Halmun (Tiyare) (Talay 2008b: 329)
le-šate-Ø
 NEG-drink.PRS-D.3MS
 ‘He will not drink’.

In C. Urmi in northwestern Iran, the negation of the habitual prefix *ci-* is *le*, e. g. *ci-šatə* ‘he drinks (habitual)’, *le-šatə* ‘he does not drink/he will not drink’ (Khan 2016, vol. 1: 286).

In the J. Urmi dialect the future particle *b-* may optionally be retained in negative clauses and is placed before the negator, e. g.

- (9) J. Urmi (Khan 2008: 85)
b-la šate-Ø
 FUT-NEG drink.PRS-D.3MS
 ‘He will not drink’.

In NENA dialects on the western periphery of the region, in the Bohtan and Cudi clusters, the negator cannot be used with an imperative form and prohibitions are expressed by combining the negator with the irrealis present, e. g.

- (10) C. Umra (Bohtan) (Hobrack 2000: 78)
la-mahk-ütən!
 NEG-speak.PRS-D.2PL
 ‘Do not speak (pl.)!’

In dialects in the eastern sector of the region the negator is used with the imperative, e. g.

- (11) J. Urmī (Khan 2008: 85)

la-⁺qtul
NEG-kill.IMP
'Do not kill!'

9. Syntax of clause linkage

9.1. Coordinating linkage

The means of coordinating linkage of main clauses in the NENA dialects of the region corresponds to the typology of the dialects of Iraq (chapter 3.4, §9.1). It may be asyndetic or marked by a particle. When marked by a particle, this is generally either (i) a clausal connective particle or (ii) a particle marking topical prominence. Examples with these two types of particle from C. Urmī are as follows:

- (12) a. *'annə miyya ⁺maxdur-e=na ⁺'allu*
They water(pl.) pour.PROG-POSS.3PL=COP.3PL on.him
'u-bərrak=əna.
and-run.PROG=COP.3PL
'They pour the water over him and run away.'
(Khan 2016, vol. 2: 401–404)
- b. *malca ⁺'avva bəxšal-u=lə baxt-u=da*
king that pound.PROG-POSS.3MS=COP.3MS wife-his=TOP
laxma marcux-u=la.
bread soften.PROG-POSS.3MS=COP.3FS
'The king pounds that (garlic) and his wife softens the bread.'
(Khan 2016, vol. 2: 405–412)

Astyndetic linkage is reported to be the norm in some dialects on the western periphery of NENA in the Bohtan cluster (Fox 2009: 110).

Turoyo links coordinated clauses either asyndetically or with a connective particle (Waltisberg 2016: 261–268).

9.2. Subordinating linkage

The general subordinating particle *d* of earlier Aramaic is widely used in Turoyo in relative clauses, clauses that are complements of verbs and clauses that are complements of adverbials:

- (13) *'i-barθo d-rhim-o-li*
the-girl REL-love.PST-D.3FS-L.1S
'the girl that I have fallen in love with'
(Jastrow 1985: 49)

- (14) *ko-ba'-no* *d-oða'-no*
 IND-want.PRS-D.1S COMP-know.PRS-D.1S
 'I want to know'
 (Waltisberg 2016: 287)
- (15) *hzeli* *d-nafq-*Ø *tarte-kurfe* *ša 'uðe* *gäläbe*
 see.PST-L.1S COMP-leave.PST-D.3MS two-snakes yellow very
šafire
 beautiful
 'I saw that two very beautiful yellow snakes came out.'
 (Waltisberg 2016: 284)
- (16) *'i-nagqa d-ko-nafq-o ...* *Fatäme ... ko-nafq-o*
 the-time COMP-IND-leave.PRS-D.3FS Fatima IND-leave.PRS-3FS
 'When it (the army) goes out, Fatima goes out.'
 (Waltisberg 2016: 302)

Asyndetic linkage is also used in these contexts, especially between a main verb and an irrealis complement clause (Waltisberg 2016: 280–283). In some cases the *d* in Turopy is replaced by subordinating particles from languages in contact such as *ki* from Kurdish and, more commonly, *'anno* and *'ənnahu* from Arabic (Waltisberg 2016: 290).

In the NENA dialects of the region the Aramaic subordinating particle *d* in relative and adverbial clauses generally has the form of a suffix *-ət* on the head item:

- (17) Ashitha (C NENA Tiyare) (Borghero 2005: 366)
yaxx-ət *'aθe-*Ø
 month-REL come.PRS-D.3MS
 'The month that is coming (i. e. next month)'
- (18) C. Urmi (Khan 2016, vol. 2: 497)
'e-⁺dan-ət *zamər-*Ø-va *'ana* *⁺šamm-ən-va*
 that-time-REL sing.PRS-D.3MS-PST I sing.PRS-D.1MS-PST
 'When he sang, I listened.'

In some dialects the original Aramaic subordinator is frequently replaced by innovative subordinating particles or particles taken from languages in contact. This process is particularly advanced in the dialects of northwestern Iran. In C. Urmi, for example, the particle *kat*, which derives historically from *ka* 'to' + *d*, originally expressing purpose, is now used to introduce direct complement clauses and relative clauses. This may be in phonetic imitation of the Kurdish subordinator *ka*:

- (19) *'e baxta kat pyay=əla laxma*
 that woman REL bake.PROG=COP.3FS bread
bədray-u=la + 'al-de-mazrak
 put.PROG-POSS.3MS=COP.3FS on-that-cushion
 ‘The woman who is baking bread puts it on the baking cushion.’
 (Khan 2016, vol. 2: 454)
- (20) *+bədday=ələ kat +'Axiqar mət-lə.*
 knows.PROG=COP.3MS that Axiqar die.PST.-L.3MS
 ‘He knows that Axiqar died.’
 (Khan 2016, vol. 2: 487)

In J. Urmi the Iranian particle *ki* frequently replaces original *d* in these contexts, e. g.

- (21) *une naše ki-la-ayol=u mar-ay-i.*
 those people REL-NEG-know.PROG=COP.3PL DEONTIC-know.PRS-D.3PL
 ‘Those people who do not know, let them know.’
 (Khan 2008: 355–359)
- (22) *mar=e ki-+ms-e brat-ox*
 say.PROG=COP.3MS COM-can.PRS-D.3MS daughter-POSS.2MS
mamqe-Ø-la.
 cause.speak-D.3MS-L.3FS
 ‘He says that he can make your daughter speak.’
 (Khan 2008: 366)

In some dialects a subordinating particle is omitted in some constructions. This is the norm across CNA and NENA dialects of the region in direct complement clauses of verbs (irrealis and realis) and relative clauses.

In most NENA dialects of the region the protasis of conditional clauses is introduced by the original Aramaic conditional particle *'ən*:

- (23) C. Urmi (Khan 2016: 516–523)
'ən- 'avə-Ø xaya, paly-ət malcuyt-i
 if-be.IRR-D.3MS alive half-GEN kingdom-POSS.1S
b-yavv-ən-na kat-ux
 FUT-give.PRS.-D1MS-L.3FS to-POSS.2MS
 ‘If he is alive, I shall give you half of my kingdom.’

In the Jewish dialects in the northwestern Iran, the conditional particle is regularly replaced by the Iranian particle *agar*:

- (24) J. Urmi (Khan 2008: 378)

agar ād mənn-i xoiš ode-t, ana
 if you from-POSS.1s request make.IRR-D.2MS I
g-de-n

IND-COME-D.1S

'If you ask me, I shall come'

Also in Turoyo, on the western periphery of the region, the Aramaic particle is replaced by a particle loaned from a contact language, either Arabic (e. g. *'ənkan*, *'iðakan*) or Kurdish (*'agar*) (Waltisberg 2016: 318–341):

- (25) Turoyo

w-iðakan 'u-zlam-ano hawxa=yo, hano latyo
 and-if the-man-this thus=COP.3MS this COP.NEG.3MS
nošo

man

'and if this man is thus, he is not a man.'

10. Concluding remarks

In the region of southeastern Turkey and northwestern Iran the major linguistic boundary in Neo-Aramaic is between the CNA and NENA subgroups at the Tigris river. CNA, especially Turoyo, is more archaic than the NENA dialects in some aspects of its phonology and morphology. Some of the NENA dialects in the Bohtan cluster on the western periphery of the NENA area exhibit more parallels with CNA than NENA dialects elsewhere. These features are likely to have arisen by convergence rather than a common ancestry, since the parallels are often only partial (e. g. the distribution of the **ā > o* shift in Ruma). Within the NENA dialects of the region there is considerable diversity. The Tiyare cluster on the whole is the most archaic in its phonology and morphology. The dialects in the eastern sector of the region are more innovative. The most innovative is the cluster of Jewish dialects in northwestern Iran.

The lexicon of the CNA and NENA dialects contain many loanwords from the languages in contact. Kurdish has had particular impact on the lexicon in all dialects. In CNA there are also numerous loanwords from Turkish and the local Arabic dialects. The Turkish influence on the lexicon intensified in the twentieth century (Tezel 2015). In the NENA dialects of northwestern Iran there are many loanwords from Azeri Turkish and Persian. These, however, appear to be a relatively recent acquisition, since the majority of them are not morphologically integrated into dialects, unlike Kurdish loanwords, which are generally morphologically adapted and represent an older lexical layer in the dialects. In the dialects of northwestern

Iran calques of Iranian and Azeri phrasal verbs containing a noun and a light verb are particularly common, e. g. C. Urmi *sohbat* 'avəd 'he converses' (literally: he makes conversation; cf. Persian *sohbat kardan*), ⁺*karavul jarəš* 'he guards' (literally: to pull guard; cf. Azer. *qarovul çəkmək*) (Khan 2016, vol. 1: 457–458).

11. Glossed texts

11.1. Turoyo

A dream (Jastrow 1985, 264–265)

The transcription and marking of prosody is taken from Jastrow's presentation of this text.

- (26) *fa b-ú-zabn-áno – kát-wa hδó ’ášm-a faṭóme*
 and at-DEM-time-this there.is-PST one.FS name-her Faṭima
kát-way-la nxiróyo ’ášm-e génč-xalil-ağa.
 there.is-PST-L.3FS lover name-POSS.3MS Genč-Xalil-Ağā
 At that time there was a woman whose name was Faṭima. She had a lover
 whose name was Genč Xalil Ağā.

- (27) *fa-háno ’í-’askar d-tərkýa mbá-lle*
 and-this DEF-army GEN-Turkey take.away.PST-L.3PL
zbót-te mahát-te b-i-’askariye ’u
 seize.PST-L.3PL place.PST-L.3PL in-DEF-army and
mbá-lle.
 take.away.PST-L.3PL
 The Turkish army took this man away. They seized him, they put him in the
 army and took him away.

- (28) *mbá-lle l-i-’askariye ’u fayíš-o*
 take.away.PST-L.3PL to-DEF-army and remain.PST-D.3FS
faṭóme táne ko-hδlm-o b-lályo
 Faṭima alone IND-dream.PRS.-D.3FS at-night
’u b-imómo b-ú-nxiroy-áyða
 and in-day about-DEF-lover-POSS.3FS
 They took him to the army and Faṭima remained alone. She dreamt night
 and day of her lover.

- (29) *yáwmo m-áy-yawme qayím-o mó̄l-la*
 day from-DEF-days stand.PST-D.3FS say.PST-L.3FS
yá-de 'ámm-o mán=yo? 'ámm-o
 VOC-mother say.PRS-D.3FS what=COP.3S say.PRS-D.3FS
b-rámšəl hzé-li hálmo 'ámm-o bárθ-i
 at-last.night see.PST-L.1s dream say.PRS-D.3FS daughter-my
'u-hálm-áyðax hòw-e xér 'amál-li
 DEF-dream-POSS.2FS be.PRS-D.3MS good say.IMP-L1s
mán-hzé-lax?
 what-see.PST-L.2FS

One day she said ‘Mother!’ She says ‘What is it?’ She says ‘Last night I dreamt a dream.’ She says ‘My daughter, your dream may be good. Tell me, what did you dream?’

- (30) *'ámm-o yá-de hzé-li 'áyno fíth-o*
 say.PRS-D.3FS VOC-mother see.PST-L1s spring open.PASS-D.3FS
b-i-dárt-áyðan 'u 'am-gab-è-d-i-çayn-áθe
 in-DET-courtyard-POSS.1PS and with-by-POSS.3MS-GEN-DET-spring-this
kát-wa dáwmo d-habúšo 'u-b-i-dawm-áθe kát-wa
 there.is-PST tree GEN-apple and-in-DET-tree-this there.is-PST
tré kárfe kóme m'álq-e 'u b-i-dawm-áθe
 two snakes black hung.PTCP-PL and in-DET-tree-this
yá-de kát-wa tré habúše 'u kát-wa
 VOC-mother there.is-PST two apples and there.is-PST
tré-kákwe me-d-i-šmayo b-i-dawm-áθe
 two-stars from-GEN-DET-sky in-DET-tree-this

She says “Mother, I saw a spring that opened up in our courtyard and by this spring there was an apple tree, and in this tree two black snakes were hanging, and in this tree, mother, there were two apples, and there were two stars from the sky in this tree.”

11.2. North-Eastern Neo-Aramaic

C. Urmî dialect

The loan of a cooking pot (Khan 2016, vol. 4, 66)

An acute accent (́) marks non-nuclear stress. A grave accent (̀) marks the nuclear stress of an intonation group. A vertical line (|) marks an intonation group boundary.

- (31) *xá-yuma +málla +Nasràdən| bərrəxš=ələ| bəškál=ələ*
 one-day mullah Nasradin go.PROG=COP.3MS take.PROG=COP.3MS

⁺*kusártā* *dénā* *mən-švàv-u.* *màr=ələ*
 cooking.pot loan from-neighbour-POSS.3MS say.PROG=COP.3MS
hál-li *xá-dana* ⁺*kusártā* ⁺*báyy-ən*
 give.IMP=L.1s one-single cooking.pot want.PRS-D.1MS
bášl-ən *jáv-o* *bušálā.* ⁺*kusártā* ⁺*júrta*
 COOK.IRR-D.1MS in-POSS.3FS stew cooking.pot big
láz-li. *báškál-o=lə*
 there.is.no-L.1s take.PROG-POSS.3FS=COP.3MS
máy-o=lə ⁺*kusártā* *bušálā* *bašūl=ələ*,
 bring.PROG-POSS.3FS=COP.3MS cooking.pot stew cook.PROG=COP.3MS
labùl-o=lə, *yáv-o=lə* *mèdra* *kà*
 take.PROG-POSS.3FS=COP.3MS give.PROG-POSS.3FS=COP.3MS again to
⁺*švàvā.* *'ína* *tré* ⁺*kusaryay-sùrə* *mattúy=əl*
 neighbour but two pots-small.PL put.PROG=COP.3MS
jáv-o. in-POSS.3FS

One day mullah Nasradin goes and takes a cooking pot as a loan from his neighbour. He says ‘Give me a pot, I want to cook stew in it. I do not have a big pot.’ He takes the pot and brings it back, cooks stew, takes it and gives it back to the neighbour, but he puts two pots in it.

- (32) *švàvā* *màr=ələ* *'áha* *tré* ⁺*kusaryatə* *kàm*
 neighbour say.PROG=COP.3MS this two cooking.pots why
muyy-é=vət? *màr-rə* ⁺*kusárt-ət*
 bring.PTCP.MS-POSS.3PL=COP.2MS say.PST-L.3MS cooking.pot-GEN
dìyyux *dàl-la* *tré* *xínə* *mènn-o.*
 GEN.2MS give.birth.PST-L.3FS two others with-POSS.3FS
yáv=əl *kàt-u* *'áv-ət* *basíma,*
 give.PROG=COP.3MS to-POSS.3MS be.IRR-D.2MS pleasant
bitáy=ələ.
 come.PROG=COP.3MS

The neighbour says ‘This is two pots, why have you brought them?’ He said ‘Your pot has given birth, two others are with it.’ He gives this to him ‘Thank you’ and comes back.

- (33) *é-šabta* *xítə* ⁺*málla* ⁺*Nasrádən* *bərráxš=əl* *mèdra.*
 that-week other mullah Nasradin go.PROG=COP.3MS again
màr=ələ ⁺*maxlèta,* *xa-*⁺*kusártā* *buš-*⁺*júrta*
 say.PROG=COP.3MS pardon one-cooking.pot more-big.FS
⁺*byáy=əvən.* ⁺*málla* ⁺*Nasrádən* ⁺*'áyn-u*
 want.PROG=COP.1MS mullah Nasradin eye-POSS.3MS
pált=əva ⁺*'al-xa-*⁺*kusart-ət* *švàvə.* *nákša*
 fall.PTCP.FS=COP.PST.3S on-one-cooking.pot-GEN neighbours plan

$\text{jríš}=\text{əva}^{\dagger}$ ${}^+ \text{bayyí}-\text{va}$ $'\text{ay}$ $\text{mayyí}-\text{va}-\text{la}^{\dagger}$
 draw.PTCP.MS=COP.PST.3S want.PRS.3MS-PST that.3FS bring.PRS-PST-L.3FS
 kà-de^{\dagger} ${}^+ \text{kusárt-u}$ $\text{yuvv-à-lə}^{\dagger}$ $\text{tré-xinə}=da$
 for-that cook.pot-POSS.3MS give.PST-D.3FS-L.3MS two-other.PL=also
 súrə mènn-o.^{\dagger} xáš-lə màr-rə^{\dagger}
 small.PL with-POSS.3FS go.PST-L.3MS say.PST-L.3MS
 ${}^+ \text{byáy}=\text{ən}$ $\text{xa-}{}^+ \text{kusárta}$ ${}^+ \text{jùrta,}^{\dagger}$ kát $'\text{zt-li}$
 want.PROG=COP.1MS one-cooking.pot big.FS COMP there.is-L1s
 $'\text{árxa}^{\dagger}$ bàsl-ən^{\dagger} xùrrac jáv-o.^{\dagger}
 guests cook.IRR-D.1MS food in-POSS.3FS

The next week mullah Nasradin goes again. He says ‘Excuse me, I want a bigger pot. The eye of mullah Nasradin had fallen on one of the pots of the neighbour. He had made a plan, since he wanted to bring back that one. For that (reason) he gave his pot as well as two others with it. He went and said ‘I want a big pot to cook food in—I have guests.

- (34) $yáv-o=lə$ $'e-{}^+ \text{kusárta}$ $buš-šapárta$
 give.PROG-POSS.3FS=COP.3MS that.FS-cooking.pot most-beautiful.FS
 ${}^+ \text{jùrta}^{\dagger}$ ká ${}^+ \text{málla}$ ${}^+ \text{Nasràdən.}^{\dagger}$ ${}^+ \text{málla}$ ${}^+ \text{Nasrádən}$
 big.FS to mullah Nasradin mullah Nasradin
 $mày-o=lə}^{\dagger}$ $mattúy-o=lə$ $bèta.}^{\dagger}$
 bring.PROG-POSS.3FS=COP.3MS put.PROG-POSS.3FS=COP.3MS house
 šváva $\text{bàxzày}=\text{ələ}^{\dagger}$ $\text{xá-yuma,}^{\dagger}$ $\text{trè-yumə,}^{\dagger}$ $\text{xá-šabta,}^{\dagger}$
 neighbour see.PROG=COP.3MS one-day two-days one-week
 $\text{trè-šabay,}^{\dagger}$ $\text{xá-yarxa,}^{\dagger}$ $\text{trè-yarxə}^{\dagger}$ ${}^+ \text{kusárta}$ lítən.^{\dagger}
 two-weeks one-month two-months cooking.pot there.is.not
 ba-dàx ví-la $'a-{}^+ \text{kusárta?}^{\dagger}$ mù ví-la
 but-how be.PST-L.3FS this-cooking.pot what be.PST-L.3FS
 ${}^+ \text{kusárta?}^{\dagger}$
 cooking.pot

He gives the biggest and most beautiful pot to mullah Nasradin. Mullah Nasradin brings it back and puts it in the house. The neighbour sees that one day, two days, one week, two weeks, one month, two months (pass) but there is no pot. ‘But what has become of the pot? What has become of the pot?’

- (35) *bərráxš=ələ* +’al-+tárr-ət +málla, | mára +málla
 go.PROG=COP.3MS to-door-GEN mullah say.PROG mullah
t-ávə-t *basíma* | +kusárta mùyy-o=vət
 FUT-BE.IRR-D.2MS well cooking.pot bring.PTCP-POSS.3FS=COP.2MS
’átən | *cmá-+dana* *kám* ’adíyya, | +bayy-àn-na. | müt
 you some-time before now want.PRS-D.1MS-L.3FS what
 +kusárta? | márrə +kusárta ’at-muyy-á-lux
 cooking.pot said.PST-L.3MS pot you-bring.PST-D.3FS-L.2MS
m-cás-lan *kat-bašl-á-t*-va xùrrac jáv-o. | ’á
 from-with-L.1PL COMP-cook.IRR-D.2MS-PST food in-POSS.3FS this
 márrə, | +kusárta, | +kusárta mät-la. | mér-rə
 say.PST-L.3MS cooking.pot cooking.pot die.PST-L.3FS say.PST-L.3MS
ba-dàx *c-óy-a* mät-la?! márrə bas-dàx
 but-how IND-BE.PRS-D.3FS die.PST-L.3FS say.PST-L.3MS but-how
c-óy-a +kusárta yàdl-a?!!
 IND-BE.PRS-D.3FS cooking.pot give.birth.IRR-D.3FS
 He goes to the door of the mullah and says ‘Mullah, if you please, I want a pot that you took some time ago.’ ‘What pot?’ He said ‘A pot that you took from me to use to cook food in.’ He said ‘Ah, the pot, the pot has died.’ He said ‘How could it have died?’ He said ‘How could a pot give birth?’

Abbreviations

Abbreviations not found in the Leipzig glossing conventions include:

D = D-suffix (i. e. direct, nominative pronominal suffix)

L = L-suffix (i. e. oblique pronominal suffix)

References

- Borghero, Roberta. 2005. The Neo-Aramaic dialect of Ashitha. Cambridge: University of Cambridge, unpublished PhD thesis.
- Duval, Rubens. 1883. *Les dialectes néo-araméens de Salamâs: textes sur l'état actuel de la perse et contes populaires publiés avec une traduction française*. Paris: F. Vieweg.
- Fassberg, Steven. 2010. *The Jewish Neo-Aramaic dialect of Challa*. Leiden & Boston: Brill.
- Fox, Samuel E. 1997. *The Neo-Aramaic dialect of Jilu*. Semitica Viva, Bd. 16. Wiesbaden: Harrassowitz.
- Fox, Samuel E. 2002. A Neo-Aramaic dialect of Bohtan. In Werner Arnold & Hartmut Bobzin (eds.), *Sprich doch mit deinen Knechten Aramäisch, wir verstehen es! 60 Beiträge zur Semitistik. Festschrift Für Otto Jastrow zum 60. Geburtstag*, 165–180. Wiesbaden: Harrassowitz.

- Fox, Samuel E. 2009. *The Neo-Aramaic dialect of Bohtan*. Piscataway, NJ: Gorgias Press.
- Fox, Samuel E. 2015. A > e raising in the Jilu dialect of Neo-Aramaic. *Aramaic Studies* 13. 161–173.
- Furman, Yulia & Sergey Loesev. 2015. Studies in the Turoyo verb. In Geoffrey Khan & Lidia Napiorkowska (eds), *Neo-Aramaic and its linguistic context*, 1–28. Piscataway, NJ: Gorgias Press.
- Garbell, Irene. 1965. *The Jewish Neo-Aramaic dialect of Persian Azerbaijan: Linguistic analysis and folkloristic texts*. Janua Linguarum. Series Practica 3. Berlin & Boston: De Gruyter Mouton.
- Heinrichs, Wolfhart. 1990. Written Turoyo. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 181–188. Harvard Semitic Studies 36. Atlanta, GA: Scholars Press.
- Hemmauer, Roland & Michael Waltisberg. 2006. Zum relationalen Verhalten der Verbalflexion im Turoyo. *Folia Linguistica Historica* 27(1–2). 19–59.
- Hoberman, Robert. 1988. The history of the Modern Aramaic pronouns and pronominal suffixes. *Journal of the American Oriental Society* 108. 557–575.
- Hobrack, Sebastian. 2000. Der neuaramäische Dialekt von Umra (Dere-Köyü). Laut- und Formenlehre. Texte. Glossar. Erlangen-Nürnberg: Friedrich-Alexander-Universität, unpublished MA thesis.
- Ishaq, Yusuf M. 1990. Turoyo: From spoken to written language. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 189–199. Harvard Semitic Studies 36. Atlanta, GA: Scholars Press.
- Jastrow, Otto. 1985. *Laut- und Formenlehre des neuaramäischen Dialekts von Mīdin im Tür 'Abdīn*. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1988. *Der neuaramäische Dialekt von Hertevin (Provinz Siirt)*. Semitica Viva 3. Wiesbaden: O. Harrassowitz.
- Jastrow, Otto. 1990. Personal and demonstrative pronouns in Central Neo-Aramaic: A comparative and diachronic discussion based on Tūrōyo and the Eastern Neo-Aramaic dialect of Hertevin. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 89–103. Atlanta: Scholars Press.
- Jastrow, Otto. 1992. *Lehrbuch der Turoyo-Sprache*. Semitica Viva 2. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1994a. *Der neuaramäische Dialekt von Mlahsō*. Semitica Viva 14. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1994b. Neuentdeckte aramäische Dialekte in der Türkei. In C. Wunsch (ed.), *XXV. Deutscher Orientalistentag. Vorträge. München, 1991*, 69–74. Deutsche Morgenländische Gesellschaft: Zeitschrift der Deutschen Morgenländischen Gesellschaft, Supplement 10. Stuttgart: Franz Steiner Verlag.
- Jastrow, Otto. 2011. Turoyo and Mlahsō. In Stefan Weninger, Geoffrey Khan, Michael Streck & Janet Watson (eds.), *The Semitic languages: An international handbook*, 697–707. Berlin & Boston: De Gruyter Mouton.
- Jastrow, Otto. 2015. Language contact as reflected in the consonant system of Turoyo. In Aaron M. Butts (ed.), *Semitic languages in contact*, 234–250. Leiden & Boston: Brill.
- Khan, Geoffrey. 2001. Quelques aspects de l'expression d'"être" en néo-araméen. In Anaïd Donabédian (ed.), *Langues de diaspora. Langues en contact*, 139–148. Faits de Langues Revue de Linguistique 18. Paris: Ophrys.

- Khan, Geoffrey. 2008. *The Jewish Neo-Aramaic Dialect of Urmî*. Gorgias Neo-Aramaic Studies. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2012. Remarks on the historical development of the copula in Neo-Aramaic. In Federico Corriente, Gregorio del Olmo Lete, Ángeles Vicente & Juan-Pablo Vita (eds.), *Dialectology of the Semitic languages. Proceedings of the IV Meeting on Comparative Semitics, Zaragoza 9–11 June, 2010*, 25–31. Sabadell: AUSA.
- Khan, Geoffrey. 2013. Phonological emphasis in North-Eastern Neo-Aramaic. In Jean Léonard & Samia Naïm (eds.), *Base articulatoire arrière: Backing and backness*, 111–132. Munich: Lincon Europa.
- Khan, Geoffrey. 2016. *The Neo-Aramaic Dialect of the Assyrian Christians of Urmî*. 4 vols. Leiden & Boston: Brill.
- Khan, Geoffrey. 2017. Ergativity in Neo-Aramaic. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *Oxford Handbook of Ergativity*, 873–899. Oxford: Oxford University Press.
- Kim, Ronald. 2008. The subgrouping of Modern Aramaic dialects reconsidered. *Journal of the American Oriental Society* 128(3). 505–531.
- Murre-van den Berg, Heleen L. 1999. *From a spoken to a written language: The introduction and development of literary Urmia Aramaic in the nineteenth century*. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Mutzafi, Hezy. 2000. The Neo-Aramaic Dialect of Maha Khtaya d-Baz: Phonology, Morphology and Texts. *Journal of Semitic Studies* 45(2). 293–322.
- Mutzafi, Hezy. 2008. Trans-Zab Jewish Neo-Aramaic. *Bulletin of SOAS* 71(3). 409–431.
- Mutzafi, Hezy. 2014. The three rhotic phonemes in Tyare Neo-Aramaic. *Aramaic Studies* 12. 168–184.
- Mutzafi, Hezy. 2015. Christian Salamas and Jewish Salmas: Two separate types of Neo-Aramaic. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 289–304. Piscataway, NJ: Gorgias Press.
- Nakano, Aki'o. 1973. *Conversational texts in Eastern Neo-Aramaic (Gzira dialect)*. Study of Languages and Cultures of Asia and Africa, A Series, no. 4. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Nöldeke, Theodor. 1868. *Grammatik der neusyrischen Sprache am Urmia-See und in Kurdistan*. Leipzig: T.O. Weigel.
- Ritter, Helmut. 1990. *Türöyö: Die Volkssprache der syrischen Christen des Tür 'Abdīn. C: Grammatik*. Stuttgart: Franz Steiner.
- Talay, Shabo. 2008a. *Die neuaramäischen Dialekte der Khabur-Assyrer in Nordostsyrien: Einführung, Phonologie und Morphologie*. Wiesbaden: Harrassowitz.
- Talay, Shabo. 2008b. The Neo-Aramaic dialects of the Tiyari Assyrians in Syria: With special consideration of their phonological characteristics. In Geoffrey Khan (ed.), *Neo-Aramaic Dialect Studies*, 39–63. Piscataway, NJ: Gorgias Press.
- Tezel, Aziz. 2015. The Turkish lexical influence on Şûrayt/Türöyo: A preliminary selection of examples. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 69–99. Piscataway, NJ: Gorgias Press.
- Tsereteli, Konstantin. 1970. Die Assyrer in der UdSSR. In Ruth Stiehl & Hans Erich Stier (eds.), *Beiträge zur alten Geschichte und deren Nachleben: Festschrift für Franz Altheim*, 2, 375–385. Berlin & Boston: De Gruyter Mouton.
- Tsereteli, Konstantin. 1977. Zur Frage der Klassifikation der neuaramäischen Dialekte. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 127. 244–253.

- Tsereteli, Konstantin. 1990. The velar spirant *g̡* in Modern East Aramaic dialects. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 35–41. Harvard Semitic Studies. Atlanta: Scholars Press.
- Waltisberg, Michael. 2016. *Syntax des Turoyo*. Semitica Viva 55. Wiesbaden: Harrassowitz.

3. Northern Iraq

3.1. Northern Iraq: overview

Geoffrey Khan and Geoffrey Haig

1. Introduction

This section treats the main languages spoken in today's northern Iraq. Politically, most of this region is part of the Kurdistan Region within Iraq, established in 2005 under a local administration, the Kurdish Regional Government, with its capital in Erbil. Historically, the region is part of what is traditionally known as Mesopotamia, geographically encompassing the lower catchment areas of the Euphrates and Tigris rivers. Across millennia, this is a region characterized by multilingualism. Four thousand years ago, the Semitic languages Akkadian and Amorite, and the genetically unclassified Hurrian and Sumerian were spoken here,¹ and today's northern Iraq is likewise multilingual, though the constellation has shifted: Semitic languages continue to be spoken in the region (Arabic and Neo-Aramaic), and while no descendants of Sumerian and Hurrian have survived, Turkic and Iranian languages have been added to the mix. Over the last century the language situation has undergone further changes due to the displacement of certain population groups. This applies in particular to the Christian and Jewish communities who speak Neo-Aramaic, a large proportion of whom now live in a diaspora outside of Iraq, and also to the non-Islamic groups of Kurdish speakers, such as the Yezidis, many of whom have been forced to leave their homeland.

Today's languages are genetically affiliated to three different language families: Semitic (Arabic, Neo-Aramaic), Iranian (Sorani and Kurmanjî Kurdish, Gorani) and Turkic (Turkman). This section contains chapters on Arabic (Procházka), Neo-Aramaic (Khan), Iranian (Haig) and Turkic (Bulut), the main spoken languages of northern Iraq. Each chapter covers the structure, and gives a survey of dialectal varieties. In keeping with the aims of the volume, we do not treat the principal official language of Iraq, Modern Standard Arabic, but focus on the (often poorly documented) local varieties. In the rest of this introduction, we touch on a number of striking similarities among the languages of this region, several of which may have arisen through language contact.

¹ For a historical perspective on the languages of Iraq see the contributions in Postgate (2007).

2. Phonology

Several dialects of the Semitic languages of the region have lost interdental consonants (*t* and *d*), generally by shifting the series to stops (*t*, *d*) or sibilants (*s*, *z*) (Procházka, chapter 3.2, §2.1.1; Khan, chapter 3.4, §5.1). This brings the consonant system closer to that of Iranian and Turkic, which do not have interdentals.

In several dialects of Sorani Kurdish a /d/ following a vowel or sonorant undergoes lenition resulting in assimilation or loss of oral articulation, the so-called ‘Zagros d’ (Haig, chapter 3.3, §3.1.1), likewise attested in several Iranian languages of western Iran (see chapter 4.1, §2.1). In Jewish Neo-Aramaic dialects spoken east of the Great Zab river (known as trans-Zab dialects) historical interdentals **t* and **d* shift to the sonorant // (Khan, chapter 3.4, §5.1). An intermediate stage of development of this appears to have been the shift of the interdentals to the voiced stop **d*. The sonorant // can then be regarded as lenition of the **d*. Such lenition, therefore, can perhaps be considered to be a convergence with that of the Sorani Kurdish ‘Zagros d’.

The assimilation of // to a stem-final /n/, which is a feature of derivational suffixes beginning with // in the Turkic dialects of the region (Bulut, chapter 3.5, §2.2) has a parallel in the Neo-Aramaic dialects, in which inflectional verbal suffixes beginning with // assimilate to a stem-final /n/ (Khan, chapter 3.4, §7.0).

Palatalization of the dorsal stops /k/ and /g/ is a feature that is found in the Kurdish, Turkic and Neo-Aramaic dialects of the region (Haig, chapter 3.3, §3.1.1; Bulut, chapter 3.5, §2.1.1; Khan, chapter 3.4, §5.1). A certain degree of areal correlation with regard to this feature can be identified in the various languages. The predominance of the palatalization of these consonants in the northern varieties of Turkic in the region, for example, corresponds to the situation in the Neo-Aramaic dialects, in which it is found mainly in dialects spoken north of Arbil (e. g. Barwar).

The existence of unaspirated voiceless stops in the consonantal system of Kurmanjî Kurdish in the north of the region (Haig, chapter 3.3, §4.1.1) has induced in NENA dialects of the area the development of unaspirated voiceless stops, which are not a historical feature of Aramaic (Khan, chapter 3.4, §5.1). The unaspirated voiceless stops themselves probably entered Northern Kurdish through Armenian (Martirosyan, chapter 2.2, §4.1; Haig, chapter 3.3).

In Turkic of the region there is a general tendency for a laryngeal /h/ to be added before initial vowels and for an initial laryngeal stop // in loanwords to be replaced by /h/ (Bulut, chapter 3.5, §2.1.1.2). This feature has spread to some Neo-Aramaic dialects in the east of the region (Khan, chapter 3.4, §5.5). Developments in the back vowels of Kurmanjî Kurdish dialects in the north of the region have been replicated in some Neo-Aramaic dialects, especially in the north-east of the region. This involves the fronting of tense back rounded vowels ([u:] > [y:]) and the raising of close-mid back vowel [o:] towards [u] (Haig, chapter 3.3, §4.1.2; Khan, chapter 3.4, §5.4).

3. Morpho-syntax

The Sorani Kurdish suffixed marker of definiteness *-eke* is used in the Jewish trans-Zab Neo-Aramaic dialects (Khan, chapter 3.4, §6.1) and in the Turkic dialects of the region (Bulut, chapter 3.5, §2.4.4.1). There is, however, a difference in grammatical integration. In the Turkic dialects it is added directly after the stem, before case and plural suffixes, as in Kurdish, whereas in Neo-Aramaic it is added after plural suffixes.

Constructions in the Kurdish noun phrase containing the genitive/attributive marker known as *ezafe* (Haig, chapter 3.3, §3.2.1.1, §4.2.1) appear to have acted as the model for developments in the Neo-Aramaic and Arabic dialects. In Neo-Aramaic the genitive marker has become an affix on the head noun of genitive constructions, in accordance with the Kurdish model, e. g. C. Barwar *bron-ət malka* son-GEN king ‘the son of the king’. In Sorani Kurdish the *ezafe* constructions consists of an invariable affix *-i* on the head noun, as in the Neo-Aramaic example just cited. In Kurmanjî, however, the *ezafe* is inflected for gender and number, and the dependent noun has oblique case-marking: *kuṛ-ē pāšā-yī* ‘son-EZ.SG.M king-OBL’. A replication of this oblique case-marking is identifiable in Neo-Aramaic in some constructions, e. g. C. Barwar *bron-ət d-o malka* son-GEN GEN-that king ‘the son of that king’. In origin the Neo-Aramaic *-ət* morpheme is a pronominal possessive suffix + attributive marker <*-eh-d. A similar construction is found in some Arabic dialects of the region, e. g. *šəkl-u l-əlwalad* appearance-his GEN-young_man ‘the appearance of the young man’ (Procházka, chapter 3.2, §2.4.2).

Several features of the verbal system of Kurdish have been replicated in Neo-Aramaic, Arabic and Turkic (Matras 2010: 75; Haig 2017). One conspicuous example of such areal features is the development of preverbal particles in Neo-Aramaic (Khan, chapter 3.4, §7.0) and Arabic (Procházka, chapter 3.2, §2.3.3). The preverbal particles in these Semitic languages appear to have a native etymology, but their distribution and function corresponds closely to the Kurdish model. This is seen most clearly in the development of preverbal particles expressing the indicative present, and particles expressing the future. There is close functional correlation between these particles in Neo-Aramaic and Kurdish. In Neo-Aramaic dialects in the Sorani-speaking area, for example, the indicative present particle is used to express also the future, as in Sorani. In the Kurmanjî-speaking area the Neo-Aramaic future particle is also used to express a habitual, as in the Kurmanjî model. There are parallels in Neo-Aramaic also to the Kurmanjî construction in which the *ezafe* particle is combined with the present indicative to express the progressive (Haig, chapter 3.3, §4.2.2), e. g. Neo-Aramaic C. Ankawa *də-k-garaš* (he is) one that pulls, i. e. ‘He is pulling’. It is noteworthy that neither Neo-Aramaic nor Arabic has replicated the Iranian subjunctive pre-verbal particle. Some dialects of Neo-Aramaic use a Turkic suffix on imperative verbs, e. g. C. Barwar *pluṭ-gən!* ‘go out (sing.)!’ (Khan, chapter 3.4, §7.0), which cor-

responds to the Turkic imperative modifying suffix {-ginen} (Bulut, chapter 3.5, §2.3.2.3).

Another innovative feature of the Neo-Aramaic and Arabic verbal systems that has developed through contact with Iranian is a perfect form that is distinct from the past perfective. Neo-Aramaic dialects exhibit various perfect forms, most of which have clear structural correlates in Kurdish. These include constructions consisting of a participle and copula (e. g. C. Barwar *grišele* ‘he has pulled’ <*griša* + *ile* pull.PTCP + COP.3MS) and constructions consisting of an invariable preverbal particle placed before the past perfective form (e. g. J. Arbel *lā grəšle* ‘he has pulled’ PARTICLE pull.PST.3MS) (Khan, chapter 3.4, §7.0). Arabic exhibits the latter type of perfect construction, e. g. *?aš kū-ğāb* what PARTICLE-bring.PST.3MS ‘What has he brought?’ (Procházka, chapter 3.2, §2.3.3). Turkic dialects in the region have also developed an innovative perfect construction consisting of a participle and copula on the Iranian model, which is in the process of replacing the old Turkic perfect containing the morpheme {-mIš} (Bulut, chapter 3.5, §2.4.2). There has also been borrowing of actual morphemes in the verbal domain. For example, the Sorani Kurdish post-verbal aspectual clitic *-ewe*, *-eve* basically meaning ‘again, back’, has been borrowed into Kurmanjî Kurdish and Neo-Aramaic.

There are innovative developments in the expression of pronominal complements of verbs in some of the languages of the region. The Turkic dialects, for example, actually borrow the pronominal clitics of Sorani, attaching them to the inflected Turkish verb to express recipients and benefactives, on the model of Iranian (Bulut, chapter 3.5, §2.4.4.2). Nominative forms of clitic pronouns are used to express the direct object in ditransitive constructions in both Arabic and Neo-Aramaic, e. g. Mosul Arabic *fatētū-k=uwwa* give.PST.1S-2MS=3MS.NOM ‘I gave it to you’ (Procházka, chapter 3.2, §2.2.2), Neo-Aramaic C. Qaraqosh *kewi-ləh=ina* give.PRS.3PL-L.3MS=3PL.NOM ‘They give them to him’ (Khan, chapter 3.4, §7.0).

All languages of the region have enclitic copulas. In the Arabic dialects these are enclitic nominative pronouns, which are a clear innovation in relation to earlier attested forms of Arabic. The Neo-Aramaic copula is likewise historically an enclitic pronoun, but in most dialects it has acquired some degree of verbal inflection due to convergence with the Iranian verbal enclitic copula. In all West Iranian languages, clause-final enclitic copulas, etymologically going back to the Old Iranian verbal copula, are present. Thus these elements are clearly an inherited feature of Iranian, and presumably provided the model for the development of overt copulas in the neighbouring Semitic languages of northern Iraq. However, the influence has gone both ways, with the Kurmanjî Kurdish dialects of the region exhibiting an additional innovative copular construction, used in e. g. clauses expressing spatial location. The copula particle in these constructions is clause-medial, and is etymologically related to the Ezafe particle, hence ultimately of pronominal origin (Haig, chapter 3.3, §4.2.2).

The alignment of Neo-Aramaic in clauses with past perfective verbs exhibits various features of ergativity, although no dialect has totally canonical ergativity. This is an innovation that has arisen by convergence with the Iranian model of ergativity, likewise restricted to clauses based on past-tense stems of the verbs. Particularly noteworthy is the manner in which verbal agreement patterns are reversed in the accusative (present, or imperfective) and ergative (past, or perfective) structures respectively. In Neo-Aramaic, agreement with the agent of transitive clauses is expressed by one set of verbal affixes in the present tenses (the D-suffixes), while the object can be expressed by another set (the L-suffixes). In the past tenses, this is reversed. For example, in the Jewish NENA dialect of Sulaymaniyah (Khan, chapter 3.4, §7.0) we have *garš-á-lu* pull.PRS-3.FS(D)-3.PL(L) ‘she pulls them’, while in the past tense (with stem-vowel alternation) we have *gərš-á-lu* ‘they pulled her’. In Sorani in the same region, verbal suffixes index the transitive subject in the present, and the object is indexed via mobile pronominal clitics, but in the past this is reversed: present tense *nā=t=nēr-īn* NEG=2SG. CL=send.PRS-1PL ‘we don’t send you’ versus past *ne=t=nerd-īn* ‘you didn’t send us’. Although the NENA and Sorani systems do not match completely (in NENA the ergative agreement patterns often extend to certain intransitives, and the argument marking in Sorani past transitives is more complex than suggested here) the overall pattern of a tense-based reversal of functions for two sets of person markers is comparable. It seems unlikely that a complex and cross-linguistically rare pattern of this nature would have arisen independently in two unrelated languages (but see also the Kartvelian “inversion”, described for Laz in Lacroix, (chapter 6.2).

Several of the Neo-Aramaic dialects of the region have developed a predominantly verb-final clausal syntax in conformity with that of Iranian and Turkic dialects. This applies in particular to the Jewish trans-Zab dialects (Khan, chapter 3.4, §8.2). All languages of the region, regardless of genetic affiliation and regardless of the position of the direct object, place the goal of a verb of movement after the verb, regardless of the position of the direct object. As a result, the Jewish trans-Zab dialects of NENA, the Iranian languages, and the Turkic languages of the region all converge on a typologically unusual combination of OV (direct object-verb) and VG (verb-goal) word order, for which Iranian languages appear to have provided the model. Similarly, all languages of the region favour finite constructions for subordination, including relativization. While this is to be expected for the Iranian and Semitic languages, the Turkic languages of northern Iraq have adapted towards the Iranian and Semitic patterns, making widespread use of finite subordination. Notably, however, the Turkic languages of northern Iraq have retained more non-finite subordination than, for example, the Turkic languages of Iran (Bulut, chapter 3.5, §2.5.1).

References

- Haig, Geoffrey. 2017. Western Asia: East Anatolia as a transition zone. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 396–423. Cambridge: Cambridge University Press.
- Matras, Yaron. 2010. Contact, convergence, and typology. In Raymond Hickey (ed.), *The Handbook of Contact Linguistics*, 66–85. Malden & Oxford: Blackwell.
- Postgate, John Nicholas (ed.). 2007. *Languages of Iraq. Ancient and modern*. Cambridge: British School of Archaeology in Iraq.

3.2. The Arabic dialects of northern Iraq

Stephan Procházka

1. Background to the language and the speech community

The region covered by this chapter comprises the Arabic-speaking parts of Iraq north of the city of Sirqāt and thus roughly corresponds to those areas of the current administrative District of Niniveh (Nīnawā) which are situated along and west of the Tigris River. The focus of the linguistic description lies on the contemporary dialect of the district's capital Mosul (al-Mawsil), which is a predominantly Arabic-speaking city.

Vernacular Arabic belongs to the Semitic branch of the Afro-Asiatic language family. Within the continuum of Arabic dialects, most varieties spoken in northern Iraq belong to the so-called Tigris branch of the Mesopotamian group. The rural Bedouin-type dialects (see below) form a distinct subgroup within Mesopotamian Arabic and exhibit many features in common with the dialects of Arabia. Some of the Arabic spoken close to the Syrian border constitutes a different layer of Bedouin Arabic: it is of the Šammar type¹ and hence also belongs to the Arabian dialect group. An overview of the main localities discussed in this chapter is provided in Fig. 1.

1.1. Brief history of the speech community

The presence of Arab tribes in pre-Islamic Mesopotamia is well-documented and can be traced back to the Parthian period and earlier. By the end of the Sassanian Empire the whole region west and south of Iraq was probably dominated by Arab pastoralists. The sedentary Arabs, whose settlements were situated mainly along the Euphrates, were culturally assimilated and often bilingual or even trilingual as they could speak Aramaic and Persian, the then dominant languages. The linguistic situation changed with the Muslim conquest of the region in the middle of the 7th century CE, which fostered massive Arab immigration from the Arabian Peninsula.² In the following centuries Arabic gradually replaced Aramaic and—to a lesser extent—Persian and Kurdish. The details of the Arabicization of the northern part of Iraq are largely unknown (Holes 2007: 123). What can be said with certainty is that it was a relatively slow process and that it only marginally

¹ Named after the large tribal confederation whose members live in northern Najd, the Syrian Desert, and parts of the Jazeera.

² For a detailed description of the Arab presence in pre- and early Islamic Iraq cf. Morony (2005: 214–253).

affected the mountainous regions east of the Tigris and the Sinjar Mountains west of Mosul. Today, in both regions Kurdish is the dominant language, with scattered enclaves of Arabic- and Neo-Aramaic-speaking small towns and villages.³

In Iraq, as well as in northeastern Syria and parts of southeastern Anatolia, two layers of vernacular Arabic can be differentiated. The older layer consists of sedentary-type dialects which are usually labeled *qəltu* dialects (after the verb *qəltu* ‘I said’) and which are mainly spoken in towns and ancient settlements. The Bedouin-type dialects (also called *gələt* dialects after the same verb) form a more recent layer because its speakers gradually infiltrated Iraq from the 13th century onwards.⁴ Thus, “for a long time the Arabic spoken in the province of Mosul consisted only of *qəltu* dialects, spoken along with varieties of Neo-Aramaic” (Abu Haidar 2004: 2, Fn.5). Until their exodus in 1950–1951, the Jews of some larger towns in Kurdistan (e. g. Arbil, ḤAqra, and Kirkūk) spoke a distinct *qəltu* dialect which constituted the only autochthonous Arabic vernacular in that region (Jastrow 1990: 5–6). Encouraged by the Arabization policy of the Baathist regime, during the 1970s and 1980s Arabic speakers from other parts of Iraq settled in Kurdish and Turkoman speaking areas of the country. But after 2003 many of them left the region and today both vernacular and standard Arabic play an ever-decreasing role in the area controlled by the Kurdistan Regional Government.

1.2. Current size, status, and religious affiliations of the speech community

No up-to-date statistical data regarding population structures of Iraq are available, but from reliable sources⁵ the number of Arabic speakers in the region can be estimated to be between 2.5 and 3 million people. As a vernacular of the official language of Iraq, the local Arabic possesses a higher status than the other languages of the region. This is particularly true for the dialect of the city of Mosul, which can be regarded as *the* prestige variety of the whole area and which constitutes without doubt the most important *qəltu* dialect with regard to number of speakers (cf. Jastrow 1979: 36).

The overwhelming majority of the Arabic-speaking population of the government are Sunni Muslims; however, some of the speakers of the *gələt* and Šammar dialects belong to Shia Islam (e. g. in the town of Rabīṭa close to the Syrian border). Christians of different confessions used to constitute as much as 15 per cent of the area’s population, but their number drastically decreased in the aftermath of the

³ Cf. the detailed map “Languages in the area of al-Mauṣil” of the *Tübinger Atlas des Vorderen Orients* (TAVO: A VIII 11).

⁴ For the underlying historical developments and their impact on the linguistic setting of Iraq cf. Holes (2007: 130–134).

⁵ Inter-Agency Information and Analysis Unit, Nineva Governorate Profile, downloaded from www.iauiraq.org/gp/print/GP-Ninewa.pdf (accessed 01 April 2013).

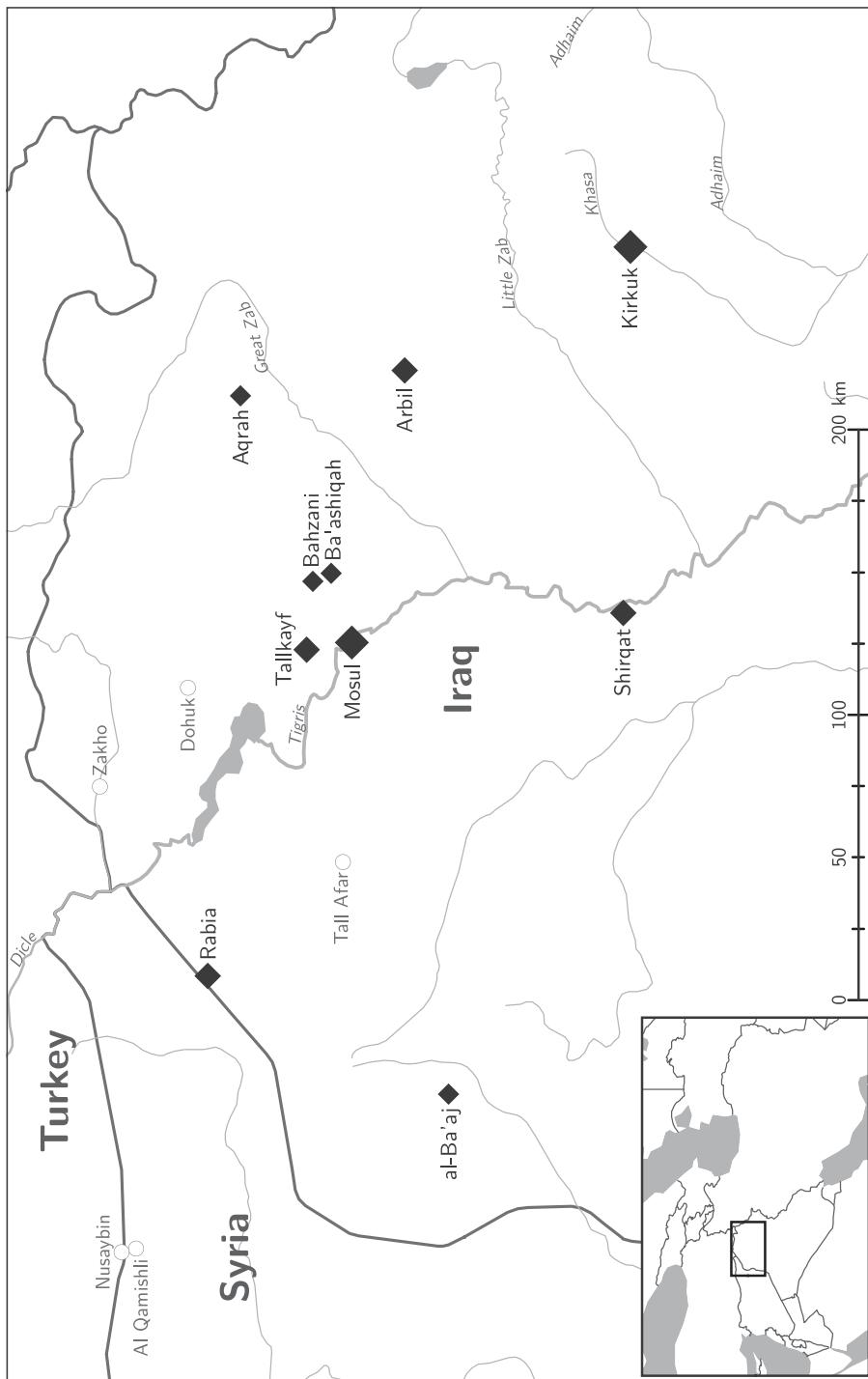


Figure 1: Locations of Arabic dialects discussed in this chapter

2003 war, which brought in its wake hostile attitudes towards non-Muslim communities.⁶ A few Arabic speakers are Yazidis, who are concentrated in the Arab villages of Bəhzāni and Baṣṣīqa, 20 km northeast of Mosul (Jastrow 1978: 24), and in the town of Tall Kayf.⁷ The conquest of large parts of northern Iraq, including Mosul, in the summer of 2014, by the ultra-Islamist forces of the so-called Islamic State has had a massive impact on the Christian and Yezidi inhabitants. Most if not all of them were killed or expelled from the regions controlled by the extremists.

1.3. Sociolinguistic situation

Due to the multi-ethnic character of the region, bilingualism of Kurdish and Arabic as well as of Turkoman and Arabic is a relatively widespread phenomenon. In most places, particularly urban settings, where Arabic speakers rarely know Kurdish, vernacular Arabic is the dominant language of communication, trade, and business. Given its status and its large number of speakers, northern Iraqi vernacular Arabic is, on the whole, definitely not endangered. However, this does not include the varieties of the displaced Christian communities, whose linguistic identity is definitely at risk because of their diaspora. The Jewish dialects of the region have already become nearly extinct, being now spoken only by a few old people, mostly living in Israel. Northern Iraqi vernacular Arabic is rarely written except in informal usage (especially by younger people in social media or text messages) or in collections of local proverbs, songs, fairy tales, and the like.

2. Linguistic description

As was mentioned, this article mainly covers the Arabic of the city of Mosul, called Maṣlāwi. But observations will also be made about characteristic features of related dialects of the region. There are no data available for the northern Iraqi *gəlat* and Šammar dialects, so they must be left out of consideration. As was pointed out by Jastrow (1979: 54–55), Maṣlāwi is a very typical variety of Iraqi Arabic in general, and of the Tigris group in particular. The city with its approximately two million inhabitants is situated in the center of the *qəltu*-speaking area and is the source of linguistic diffusion into its surroundings. As there are only

⁶ Most of the area's Christians left Iraq for Europe and America; but tens of thousands fled to politically stable Iraqi Kurdistan (source: webpage of the German NGO "Gesellschaft für bedrohte Völker" www.gfbv.de/pressemit.php?id=2958&stayInsideTree=1 (accessed 26 March 2013).

⁷ ar.wikipedia.org/wiki/%D8%AA%D9%84%D9%83%D9%8A%D9%81 (accessed 25 March 2016).

minor differences between Muslim and Christian Maṣlāwi speakers,⁸ we have not indicated the religion of the speakers of the examples given below.

2.1. Phonology

2.1.1. Consonants

The following table includes the **consonantal phonemes** of Mosul Arabic.

Table 1: The consonantal phonemes of Mosul Arabic

	Bilabial	Labio-dental	Inter-dental	Dental-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive									
voiceless	p			t t [t̪]		k	q		(?)
voiced	b			d		g			
Affricate									
voiceless					č [tʃ̪]				
voiced					ğ [dʒ̪]				
Fricative									
voiceless	f	t̪ [θ̪]	s̪ [s̪]	š̪ [ʃ̪]	x	h̪ [h̪]			h
voiced		d̪ [ð̪]	z̪	g̪ [ɣ̪]		ç̪			
Nasal	m		n						
Lateral			l̪ [l̪]						
Trill				(r)					
Approximant	w				y				

A salient feature of Mosul Arabic and of (almost all) other Tigris branch dialects is that /r/ is only a marginal phoneme since in most words the two Old Arabic (= OA) phonemes /r/ and /ğ/ have merged into /ğ/⁹ (Jastrow 1979: 38), e. g. *baqağ* ‘cows’ (<*baqar*), *ğəğəl* ‘foot’ (<*riğl*), *dağəb* ‘path’ (<*darb*).¹⁰ In the sequence ڻğC, /ğ/

⁸ Cf. Jastrow 2004: 142; this was confirmed by my informants. The dialect of the Jews, however, significantly deviated from Muslim and Christian Maṣlāwi.

⁹ The sound /r/ is often found instead of /ğ/ in numerals. It is also heard in some lexicalized words like *rəkəb* ‘he mounted (a horse)’ and *farrag* ‘he emptied’ (Jastrow 1979: 38).

¹⁰ Examples, for which no source is given, are from own research with informants and unpublished BA/MA-theses from the University of Vienna, particularly by Hiam Wadie and Lea Bäumler.

is frequently elided, resulting in the lengthening of the preceding vowel: e. g. *qōṣa* (< *qəḡṣa*) ‘loaf (of bread)’. The sound shift *r* > *g* is an ancient Iraqi feature and discussed in detail by Blanc (1964: 20–25).

In the dialects spoken in the western part of northern Iraq, /r/ and /g/ are independent phonemes: *Rabīṭa rəħtu* ‘I went’; *?arnab* ‘rabbit’ (Abu Haidar 2004: 5); *al-Baʕāğ¹¹* *rəğəl* ‘foot’ (Talay 1999: 68). The same was true for all the Jewish dialects of Kurdistan (Jastrow 1990: 6).

Most dialects of the region possess the three interdental fricatives /d/, /t/, /d̪/. Exceptions are Bəħzāni, where they have shifted to sibilants (Jastrow 1978: 34–38), and Rabīṭa where they have become dental stops: e. g. *hāda* ‘this’, *mətəl* ‘like’, *hadq* ‘luck, fate’ (Abu Haidar 2004: 4). Concerning the sounds /p/, /g/, and /č/, see below, section 3. A common feature of Mosul Arabic is the devoicing of consonants in word-final position; this applies particularly to /b/ and /d/, but not to fricatives such as /f/ and /ɣ/ (Jastrow 1979: 41). Abu Haidar (2004: 6) reports examples in the dialect of Rabīṭa of the elision of /f/: for instance *mōraf* ‘I don’t know’ (< *mā ařraf*).

2.1.2. Vowels

Mosul Arabic has two short (/ə/ and /a/) and five long (/ā/, /ē/, /ī/, /ō/, /ū/) vowel phonemes (Jastrow 1979: 39). Diphthongs are monophthongized in Mosul but occur in Rabīṭa: *mawt* ‘death’, *bayt* ‘house’ (Abu Haidar 2004: 9). Frequently found are the shifts *ū* > *ō* and *ī* > *ē* when the vowel precedes a velarized or back consonant: *sōga* ‘picture’ (< *siura*); *batṭexa* ‘sugar melon’ (< *batṭixa*).

Diachronically seen, in many words an original vowel /ā/ has shifted to /ē/ or to /ī/ under the influence of /i/ and /ī/, respectively, in the same word. This phonological process (called *?imāla* by the Arab grammarians) is a characteristic of all *qəltu*-dialects. Examples: *ğawēməf* ‘mosques’ (< *ğawāmi*); *faniğīn* ‘cups’ (< *fanāğīn*)¹² (Jastrow 1979: 40). In Mosul Arabic this sound shift also happens when a suffix that contains a long or short /i/ is attached to a noun: e. g. *bəstān* ‘garden’, *bəstēn-či* ‘gardener’ (Blanc 1964: 47).

The feminine suffix has two allomorphs, *-i* and *-a*, which depend on the preceding consonant (Jastrow 1979: 40): The allomorph *-a* appears after velarized and back consonants, and *-i* in all other cases. Examples are: *bēḍa* ‘egg’, *mēlha* ‘salty’, *fāġa* ‘mouse’, *ğeġi* ‘hen’, *lēli* ‘night’, and *ğāmūsi* ‘female buffalo’.

¹¹ The *qəltu* Arabic speaking inhabitants of the town are of the Xawētna tribe whose dialect belongs to the Euphrates branch.

¹² In Rabīṭa only the shift *ā* > *ē* occurs: e. g. *dakēkīn* ‘shops’ (and not **dakīkīn*; Abu Haidar 2004: 9).

2.1.3. Syllable structure

The possible syllables are C_v, C_{v̄}, C_{v̄}C, and C_{v̄}C. The syllable CC_{v̄}- is usually realized with the epenthetic, non-phonemic vowel /ə/, which traditionally has been transcribed as ^ə: e. g. ^ənǵōh ^ənzūr xālna ‘we are going to visit our uncle’.¹³ The sequence C_vCC#¹⁴ is always changed to C_vCəC# (Jastrow 1979: 41), e. g. *d̥ħaġ* ‘noon’, *mələh* ‘salt’, *qəġən* ‘horn’ (but *qəġnu* ‘his horn’).

If, due to inflection or the attaching of suffixes, a syllable of the type Cə becomes unstressed, the vowel is in most cases elided: *šáġəb* ‘he drank’, **šáġábtu* > ^ə*šáġəbtu* ‘I drank’. A typical feature of the *qəltu* dialects is that, if a pronominal suffix is attached to a noun or verb, elision of the final vowel is prevented by moving the stress onto the word’s last syllable: *mašákəl* ‘problems’ – *mašákələm* ‘their problems’; *qəssə* ‘story’ – *qəssətī* ‘my story’; *yāxəd* ‘he takes’ – *yāxádu* ‘he takes it’ (in contrast to *yāxdu* ‘they take’).

2.2. Nominal morphology

2.2.1. Inflection

There are no case markers in northern Iraqi vernacular Arabic. There is masculine and feminine gender, the latter usually marked by the morpheme *-a/-i*. Nouns can generally be inflected according to the three numbers singular, dual, and plural: the singular is not marked; the suffix of the dual is *-ēn*; and the plural is formed by suffixes (in particular *-āt*) or by the internal reshaping of the word pattern.

Table 2: Number inflection on nouns

	time, instance	door	path
SG	<i>kaġġa</i>	<i>bāb</i>	<i>daġəb</i>
DUAL	<i>kaġġtēn</i>	<i>bābēn</i>	<i>daġbēn</i>
PL	<i>kaġġāt</i>	^ə <i>bwāb</i>	^ə <i>dḡūb</i>

In Rabī‘a and Mosul the unit nouns of collectives often have the ending *-āyi*: e. g. *nəġmāyi* ‘a star’, *kəbbāyi* ‘a patty of crushed wheat and minced lamb’, *bayḍāyi* ‘an

¹³ In the former Jewish dialects of Kurdistan an initial CC- was realized as CəC-. In certain verbal patterns it also occurred without an epenthetic vowel: e. g. *ṣrab* ‘drink!’. Examples of CəC are *zəbōf* ‘week’, *ṣərəbtu* ‘I drank’, and *nənām* ‘we sleep’ (Jastrow 1990: 32).

¹⁴ If the two consonants are identical, only one is pronounced: C₁vC₂C₂ > C₁vC₂. Only in the dialect of the village of Bəħzāni syllables ending on -CC# are found (Jastrow 1979: 41).

egg' (Abu Haidar 2004: 10). In Mosul most of these forms are used besides those ending in *-i/-a*; thus one also hears *nəğmi* star' and *bēğda* 'egg'.

2.2.2. Pronouns

The following table shows the independent and bound personal pronouns of Mosul Arabic, based on Procházka (2006–2007: 125) and Jastrow (1979: 42–43).

Table 3: Pronouns (independent and bound) of Mosul Arabic

	independent forms	bound forms after C — after V
1 SG	<i>ʔana</i>	<i>-i/-ni</i> — <i>-yi/-ni</i>
2 SG M	<i>ʔənta</i>	<i>-ak</i> — <i>-k</i>
2 SG F	<i>ʔənti</i>	<i>-ki</i> — <i>-ki</i>
3 SG M	<i>huwwa (hīnu)</i>	<i>-u</i> — <i>-nu</i>
3 SG F	<i>hīya</i>	<i>-a</i> — <i>-ha</i>
1 PL	<i>nəħna</i>	<i>-na</i> — <i>-na</i>
2 PL	<i>ʔəntəm</i>	<i>-kəm</i> — <i>-kəm</i>
3 PL	<i>hīyəm</i>	<i>-əm</i> ¹²⁷ — <i>-ħəm</i>

The form *hīnu* for 3M.SG seems to have become obsolete as only *huwwa* is heard in contemporary speech (cf. also Jastrow 1979: 42). The independent forms are used mainly in the subject position, but they (or shortened variants of them) can also indicate a second pronominalized direct object: *satētū-k-uwwa* 'I gave it (masc.) to you', *satētū-k-īya* 'I gave it (fem.) to you' (Jastrow 1979: 43); *mā-nəħtī-ki-hīyəm* 'we do not give them to you' (Jastrow 2004: 143/8); *sallamītū-ləm hīyəm* 'I handed them over to them' (Jastrow 2004: 143/7).

2.3. Verbal morphology

2.3.1. Inflection

There is a twofold system for the inflection of finite verbs: the suffix-based conjugation for the **perfect**, and the mostly prefix-based conjugation for the **imperfect**. The inflectional affixes are attached to the perfect and imperfect stem, respectively. Two sets of vowel patterns exist for the basic stem (also called Form I): *CaCaC* and *CəCəC* for the perfect, and *-CCaC* and *-CCəC* for the imperfect.

¹⁵ In Rabi'a *-ən* is used (as in Anatolia): *baytən* 'their house' (Abu Haidar 2004: 11).

Table 4: Perfect of *šagab* ‘to drink’

	SG	PL
1	² šgəb-tu	² šgəb-na
2 M	² šgəb-ət	² šgəb-təm
2 F	² šgəb-ti	
3 M	šagab	šagb-u
3 F	šagb-ət	

Table 5: Imperfect of *šagab* ‘to drink’

	SG	PL
1	ʔa-šgab	nə-šgab
2 M	tə-šgab	tə-šgab-ūn
2 F	tə-šgab-īn	
3 M	yə-šgab	yə-šgab-ūn
3 F	tə-šgab	

The final *-n* of *-īn* and *-ūn* is dropped when a pronominal suffix is attached: e. g. *yəqšaīūn* ‘they see’, but *yəqšaīū-k* ‘they see you (M.SG)’.

Table 6: Imperative of *šagab* ‘to drink’

	SG	PL
M	² šgab	² šgab-u
F	² šgab-i	

There exists an enclitic particle *=ze*¹⁶ which is used to emphasize imperative forms: e. g. *qūmī-ze* ‘Stand up (F.SG)!’.

2.3.2. Derivation

The derivation of verbs is limited to nine forms for triconsonantal roots and two for quadricsonantal roots. Each form has a set of patterns for the perfect and imperfect base as well as for the active and passive participle.

¹⁶ This appears to parallel the use of a clitic particle *īš* in Iraq-Turkic, which may be added to finite verbs (including imperatives) to impart a sense of emphasis. In Iraq-Turkic, it can be plausibly attributed to borrowing of the Kurdish clitic particle *=īš* / *=zī*; cf. Bulut, this volume, chapter 3.5, §3.4.3.

2.3.3. Verbal modifiers

A set of verbal modifiers is used in Maṣlāwi Arabic and other sedentary dialects of the region¹⁷ to express future tense as well as the durative, habitual, perfect, and progressive. The last is expressed by the prefix *qa-* (*qad-* with 1SG):¹⁸ e. g. *qa-yfat-tlūnu b-əs-sōq* ‘they are showing him around the market’ (Jastrow 1979: 47); *ən-nəswān qa-yəxbəzūn xəbəz* ‘the women are baking bread’. Progressive present tense is also indicated by the invariable prefix *kē-*, e. g. *gēħət əl-pāča kē-təṭlař* ‘the smell of the *pāča* is ascending’.¹⁹

In Mosul continuous or habitual past tense is indicated by inflected forms of *kān* ‘to be’ followed by an imperfect: e. g. *kənna nəxbəz* ‘we used to bake bread ~ we were baking bread’ (Jastrow 1979: 47). In Rabī'a, the same is expressed by the prefix *ka-*: e. g. *ka-təħləb ən-nařğēt* ‘she used to milk the ewes’ (Abu Haidar 2004: 10).

The present perfect aspect is expressed by the invariable proclitic *kən-* followed by a verb in the suffix conjugation: e. g. *w-nəħna kən-ħaġanna l-ħaġīn w haddaqnānu* ‘and we have already kneaded the dough and have prepared it’ (Jastrow 1979: 72/34); *mā kən-ȝšəřt əš-šōbak?* ‘Haven’t you seen the rolling pin yet?’ (Jastrow 1979: 70/29).

In the dialects of the Jews of Kurdistan, the verbal modifier *kū* was used for both present and past time events: *?aš kū-trīd?* ‘What do you want?'; *?aš kū-ȝāb* ‘What has he brought?’ (Jastrow 1990: 63–67).

Future tense is indicated by the invariable *ȝayyəħ*: e. g. *?aš ȝayyəħ təħmalūn ġada b-əl-lēl?* ‘What will you (PL) do tomorrow evening?’; *ȝayyəħ təzzawwaġ* ‘she will marry’. Jastrow (1979: 47) mentions two other modifiers, *də-* and *kawi* (*də-*): e. g. *balki də-yəbħat* ‘maybe he will send’; *kawi ?aġōħ ~ kawi d-aġōħ* ‘I shall go’. In the village of Bəħzāni *tə-* is found: e. g. *tīqūl* (< **tə-yqūl*) ‘he will say’ (Jastrow 1978: 303).

2.4. Syntax

2.4.1. The syntax of copula clauses

The only known dialects of northern Iraq which use a copula in expressions “A is B” are all spoken east of the Tigris. Among them is the dialect of the mainly Yazidi

¹⁷ An exception is the dialect of the Xawētna – probably because of Bedouin influence (Talay 1999: 178).

¹⁸ The prefix goes back to the participle *qāħid* ‘sitting’.

¹⁹ This modifier is not found in Jastrow’s texts, but in the texts published by Socin 1882–83 (see also Jastrow 1979: 47, Fn.33). According to my informants it is typical of Christian Maṣlāwi speakers.

village of Bəhzāni and the former Jewish dialects, which also used an enclitic copula.

- (1) Bəhzāni (for details see Jastrow 1978: 135)
báyt-a-hu
 house-3SG.F-COP.3SG.M
 ‘It is her house.’
- (2) Kurdistan Jews (Jastrow 1990: 37 f)
b-əl-bēt-nəħne
 IN-DEF-house-COP.1PL
 ‘We are at home.’

2.4.2. Nominal attribution

There are different constructions for nominal attribution in the dialect of Mosul. One possibility is direct asyndetic attribution which requires the construct state of the head noun that is morphologically marked only with feminine nouns (examples 3a and b). A less common way of expressing the same meaning is by an anticipatory pronoun plus the preposition *l-* ‘for’ (example 3c).

- (3) Mosul (Jastrow 2004: 142/5)
 - a. *sīyār-t* *əl-pōlīs* *ğə-tt* *falē-na*
 car-F.CONSTR.ST DEF-police come-PF.3SG.F to-1PL
 ‘The police car came to us.’
 Mosul (own data)
 - b. *šəkl* *əl-walad*
 appearance.CONSTR.ST DEF-young_man
 ‘the young man’s outer appearance’
 Mosul (Jastrow 1979: 49)
 - c. *šəkl-u* *l-əl-walad*
 appearance-3SG.M for-DEF-young_man
 ‘the young man’s outer appearance’

Extremely frequent is analytical attribution by means of the so-called genitive marker *māl* or *mālət*. In this case the whole expression can be definite even if the head noun is morphologically marked as indefinite.²⁰ (4) and (5) are two examples of definite expressions with an indefinite head noun, whereas (6) and (7) are examples of expressions with definite head nouns.

- (4) Mosul (Jastrow 2004: 143/7)

²⁰ There is evidence for an analogous usage in the Christian Baghdadi dialect (Blanc 1964: 125–126) and in the Jewish dialect of Hīt (Khan 1997: 93).

?axad-u ?aqwāl mālət-na
 take-PF.3.PL testimony.PL GEN.MARKER-1PL
 'They recorded our testimonies.'

- (5) Mosul (Jastrow 2004: 143/7)
daftar xədma māl zōg-i
 notebook service GEN.MARKER husband-1SG
 'my husband's service record'

- (6) Mosul (Jastrow 2004: 143/18)
asbāb əl-xurūğ mālət-i mən əl-ṣirāq
 reason.PL DEF-exit GEN.MARKER-1SG from Iraq
 'my reasons for leaving Iraq'

- (7) Mosul (Jastrow 2004: 144/33)
əl-kəg̪ši māl-u
 DEF-paunch GEN.MARKER-3SG.M
 'his paunch'

An indefinite expression with an indefinite head noun is:

- (8) Mosul (Jastrow 2004: 144/28)
ʃdām māl ḡanam
 bone.PL GEN.MARKER sheep
 'bones of sheep'

2.4.3. Direct objects

Besides asyndetically expressed direct objects, a second way to mark a direct object if the latter is a definite noun is by linking the noun to the predicate by the preposition *l-*²¹ preceded by an anticipatory pronoun that agrees in number and gender with the object. This construction, which is found in several other Semitic languages (Khan 1984), is commonly regarded to be an Aramaic substratum in (northern) Iraqi Arabic (Blanc 1964:130; Contini 1999: 105; Weninger 2011: 750). In the dialect of Mosul it is very frequent and often in free variation with constructions without marking the direct object:²²

- (9) Mosul (own data)
 a. *əmbēḥa qšəf-tū-nu l-abū-yi*
 yesterday see-PF.1SG-3SG.M for-father-1SG

²¹ The preposition is sometimes omitted (Jastrow 1979: 49).

²² There are many instances of it in the texts which Jastrow 1979 recorded from a Muslim woman. In Jastrow 2004, recorded from a Christian woman, this construction is much rarer.

or:⁷

- b. *'mbēħa qšəf-tu ɻabū-yi*
 yesterday see-PF.1SG father-1SG
 ‘Yesterday I saw my father.’

2.4.4. Negation

Several particles are used in different situations for negation. They all immediately precede the word to be negated. The particle *mā* is the most frequent way to negate finite verbs, both perfect and imperfect.

- (10) Mosul (Abu-Haidar 2002: 6)

- ħamāt-i kəll ġəmġ-a mā dāg-ət-l-i*
 mother-in-law-1SG all life-3SG.F NEG turn-PF.3SG.F-to-1SG
bēl
 attention
 ‘All her life my mother-in-law didn’t pay any attention to me.’

- (11) Mosul (Abu-Haidar 2002: 7)

- kaġġt-ēn sallam-tu ġalayy-a w mā sallam-ət ġalayy-i*
 time-DUAL greet-PF.1SG on-3SG.F and NEG greet-PF.3SG.F on-1SG
 ‘Twice I greeted her and she didn’t greet me.’

- (12) Mosul (Abu-Haidar 2002: 7)

- mā n-ṭīq nə-dfa' əl-īgāg*
 NEG IPF.1PL-can IPF.1PL-pay DEF-rent
 ‘We cannot pay the rent.’

Occasionally *mā* is also used to negate active or passive participles:

- (13) Mosul (Abu-Haidar 2002: 7)

- ṛa-ṣūf mā lēbəs qāṭ*
 IPF.1SG-see NEG wear.AP.M.SG suit
 ‘I see you are not wearing a suit.’

The particle *lā* negates the imperative, as in (14), and occasionally also an optative (15). When followed by *kān* it expresses the speaker’s wish that an event should not have happened, as in (16).

- (14) Mosul (Abu-Haidar 2002: 3)

- lā tə-nṭəy-əm ahamməyyi ha-l-qadd!*
 NEG IPF.2.M.SG-give-3PL attention DEM-DEF-extent
 ‘Don’t give them so much importance!’

- (15) Mosul (Abu-Haidar 2002: 5)

lā yə-qšaʕ-ōn yōm gāha
 NEG IPF.3-see-PL day rest
 ‘May they never see a day of rest!’

- (16) Mosul (Abu-Haidar 2002: 5)

lā kān sāhaǵ-təm l-nəs̪s̪ əl-lēl
 NEG be.PF.3SG.M stay.awake-PF.2.PL to-middle DEF-night
 ‘You shouldn’t have stayed up until midnight!’

The particle *mū* negates a non-verb (17). Sometimes, in admonitory speech, it precedes an imperfect, shown in (18):

- (17) Mosul (Abu-Haidar 2002: 8)

mū əsm-a salwa maǵt axū-nu
 NEG name-3SG.F Salwa wife brother-3SG.M
 ‘His brother’s wife is not called Salwa.’

- (18) Mosul (Abu-Haidar 2002: 8)

mū tə-ǵ-ōn ʃəbʃān-īn
 NEG IPF.2-come-PL replete-PL
 ‘You (PL) must not arrive feeling full!’

2.4.5. Syntax of clause linkage

Coordinative clauses are mostly linked by conjunctions, particularly *w* ‘and’, *bass*/ *lākən* ‘but’, and *wəlla* ‘or’. Subordinate clauses which are dependent on verbs of speaking, of belief, and the like are very often not introduced by a conjunction, shown in (19–20). Less frequently, the same may be true of final clauses, as in (21).

- (19) Mosul

qāl mā yə-ǵi
 say.PF.3SG.M NEG IPF.3SG.M-come
 ‘He said that he will not come.’

- (20) Mosul

mā n-ṣaddəq ǵayyəh yə-zzawwaǵ-ūn
 NEG IPF.1PL-believe FUT IPF.3-marry-PL
 ‘We don’t believe that they will marry.’

- (21) Mosul (Jastrow 1979: 70/33)

ləmmən-mā nə-ǵi nə-ʃləq-u
 when-PRTCL IPF.1PL-come IPF.1PL-light-3SG.M
 ‘When we come to light it.’

Final clauses are frequently marked by the verbal modifier *də-*.

- (22) Mosul (Jastrow 1979: 75/42)

w yə-ği yə-ğlī-nu yə-şgab-u
 and IPF.3SG.M-come IPF.3SG.M-boil-3SG.M IPF.3SG.M-drink-3SG.M
də-yə-stağīh
 MOD-IPF.3SG.M-recover
 ‘To recover, he boiled it and (then) drank it.’

- (23) Mosul

ğā bağād sani də-y-axəd həşşət-u
 PF.3SG.M.come after year MOD-IPF.3SG.M-take share-3SG.M
 ‘After one year he came to take his share.’

Conjunctions are frequently found in temporal and causal clauses as well as in conditionals. Many temporal conjunctions consist of a preposition followed by a particle *mā*. Numbers (24–29) are examples of subordinated temporal clauses. An example of a temporal clause without a subordinating conjunction is (30).

- (24) Mosul (Jastrow 1979: 70/33)

nə-şəl əb-batn-u hatta y-şīg ahmag
 IPF.1PL-light in-belly-3SG.M until IPF.3SG.M-become red
 ‘We heat its inner part until it gets red-hot.’

- (25) Mosul (Jastrow 1979: 72/35)

n-şəff-u ... ələmmən-mā nə-nzal lī-ğawwa
 IPF.1PL-line.up-3SG.M until-PRTCL IPF.1PL-descend to-below
 ‘We put them in a row ... until we reach the bottom line.’

- (26) Mosul

ləmmən sāğ əd-dəhəğ gāh fənd
 until PF.3SG.M-become DEF-noon PF.3SG.M-go at
əl-kabābči
 DEF-kebab_seller
 ‘When it was noon he went to the kebab seller.’

- (27) Mosul (Jastrow 1979: 72/39)

bağd-mā yə-stawi y-ğib-ūn tāhīn
 after-PRTCL IPF.3SG.M-be.cooked IPF.3-bring-PL flour
 ‘After it has fully cooked they bring flour.’

(28) Mosul

<i>ʃaqəb-mā</i>	<i>yə-stawi</i>	<i>y-qūm-ūn</i>
after-PRTCL	IPF.3SG.M-be.cooked	IPF.3-stand_up-PL
<i>y-hət̪-ū-l-u</i>	<i>māy</i>	
IPF.3-put-PL-for-3SG.M	water	

‘After it has fully cooked, they put water into it.’

(29) Mosul

<i>qabəl-mā</i>	<i>y-ğōh</i>	<i>ʃa-l-bēt</i>	<i>əʃtagā</i>	<i>təffāh</i>
before-PRTCL	IPF.3SG.M-go	to-DEF-house	buy.PF.3SG.M	apples.COLL

‘Before he went home he bought some apples’

(30) Mosul (Jastrow 1979: 68/24)

<i>w tə-ği</i>	<i>s-samak-i</i>	<i>w</i>	<i>y-ğəğğ-ū-ha</i>	<i>y-şīd-ū-ha</i>
and IPF.3SG.F-come	DEF-fish-F	and	IPF.3-pull.out-PL-3SG.F	IPF.3-hunt-PL-3SG.F

‘When the fish comes, they pull it out and catch it.’

The following is an example of a causal clause.

(31) Mosul

<i>mā</i>	<i>ʃtagē-tu</i>	<i>ṭamāṭa</i>	<i>mbūg</i>	<i>kān-ət</i>	<i>əktīg</i>
NEG	buy-PF.1SG	tomatoes.COLL	because	be-PF.3SG.F	much
<i>gālīy-i</i>					expensive-F

‘I did not buy tomatoes because they were too expensive.’

Conditional clauses are introduced either by *?ida* (32) or by *lō* (33).

(32) Mosul (Jastrow 1979: 70/28)

<i>yə-tmāsak</i>	<i>aktağ</i>	<i>ida</i>	<i>ṣtabar-na</i>	<i>ʃalē-nu</i>
IPF.3SG.M-cohere	much.ELA	if	be.patient-PF.1PL	on-3sg.m
<i>šwayya</i>				little

‘It (the dough) coheres better if we are a little bit patient with it.’

(33) Mosul

<i>lō</i>	<i>qšəf-tī-nu</i>	<i>salləm-ī-l-i</i>	<i>ʃalē-nu</i>
if	see-PF.2.F.SG-3SG.M	greet-IMP.2.F.SG-for-1SG	on-3SG.M

‘If you (F.SG) see him, give him my regards!’

Relative clauses follow the head noun and are introduced by the particle *əlli* if the head noun is definite.

(34) Mosul

<i>ʔn-ğōḥ</i>	<i>ma 'ā-hām</i>	<i>'ānd</i>	<i>xālāt-na</i>	<i>əlli</i>	<i>tə-skən</i>
IPF.1PL-go	with-3.PL	at	aunt-1PL	REL	IPF.3SG.F-live
<i>b-²mħall-ət</i>		<i>əl-ğūlāq</i>			

in-quarter-F.CONSTR.ST toponym

‘We go with them to our aunt, who lives in the Ğūlāq neighborhood.’

3. Additional notes on contact-phenomena

We may assume that linguistic interaction of speakers of Kurdish, of Turkoman (and, until the 1920s, of Ottoman Turkish), and of Bedouin and sedentary Arabic is a matter of everyday life in most of the towns and villages of the region in question. As far as vernacular Arabic is concerned, these contacts have mainly influenced the lexicon, particularly terms of material culture, such as food, housewares, tools, and the like.²³

A direct result of these borrowings is the establishment of sounds in the vernacular which were originally unknown to Arabic, particularly /p/, /g/, and /č/. Examples of words with these sounds are *pāča* ‘stew made of the head, feet, and neck of an animal’, *parda* ‘curtain’, *lagan* ‘metal wash basin’, *glāš* ‘glass’, *šūč* ‘fault’, *čāy* ‘tea’.²⁴

In the dialect of the Iraqi Xawētna (who mainly reside in al-BaYYāğ), the sounds /č/ and /g/ are not only found in foreign loans, but also due to the influence of the Bedouin dialects, where they go back to the OA consonants *k* and *q* (Talay 1999: 22, 31). Talay (1999: 19, 22) emphasizes the strong Bedouin element in Southern Xawētna Arabic, which certainly includes its Iraqi branch.

We may assume that calques are numerous and that there is a good deal of mutual influence in phraseology. A prominent example is the phrase ‘What’s the news?’ (35), which is an exact calque of its Turkish equivalent.²⁵ In-depth studies on the lexicon of the languages involved could certainly detect many other such loan translations.

(35) a. Mosul

<i>ʔaš</i>	<i>aku</i>	<i>mā-ku?</i>
Q.what	there.is	there.is.not

²³ Examples can be found in the texts published by Jastrow 1979 and 2004 as well as in the dictionary by Bakrī 1972.

²⁴ As can be seen from the examples, these sounds are also found in loanwords taken from languages from outside the region, particularly English.

²⁵ Jastrow (1990: 13) states that the Turkish of the Ottoman culture had more influence on the Jewish Arabic dialects than did the local Azeri (also known as Turkoman, see Bulut, this volume, chapter 3.5).

b. Turkish

<i>ne</i>	<i>var</i>	<i>ne</i>	<i>yok</i>
Q.what	there.is	Q.what	there.is.not
'What's the news?'			

The influence of Iranian and Turkic languages on the morphology²⁶ and syntax²⁷ of Northern Iraqi Arabic seems to be marginal or even non-existent. One reason for this might be that—at least in cities like Mosul—Arabic speakers are rarely bilingual, a fact which prevents far-reaching influence on morphological and syntactical structures. There are, however, no extended corpora of recordings of unmonitored everyday conversations and it is entirely possible that such texts would reveal a good deal more contact phenomena than has hitherto been discerned.

A possible influence of Kurdish in the dialect of the village of Bəhzāni²⁸ is the shift of the three interdental fricatives /d/, /t/, /d̪/ to the sibilants /z/, /s/, /z̪/, which resembles very much the substitution of Arabic loanwords in Kurdish (and Turkish). Whether or not the analogous shift to dental stops in Rabī'a is the result of a contact with Syrian sedentary dialects is hard to say. The town is very close to the Syrian border, though all adjacent dialects have interdentals. Hence one would have to assume that this shift has been caused by one of the urban Syrian prestige varieties (Damascus, Aleppo). This is not impossible as in the very same dialect the typical Syrian verbal modifier *b-* is also found: e. g. *b-tākəl* ‘you eat’ (Abu Haidar 2004: 10).

4. Glossed text sample

The text was provided by a male speaker of Mosul Arabic, born approximately in 1942. The recording was made by Hiam Wadie in May 2004, who also supplied an initial rough gloss, later modified for the present purposes.

- (36) *ğā* *š-šēṭān* *ɻa=l-maṣlāwi*
 come.PF.3M.SG DEF-devil on=DEF-Maṣlāwi
 w=qal=l-u
 and=say.PF.3SG.M=TO-3SG.M
 ‘The devil came to the Maṣlāwi and said to him,’

²⁶ An example of influence on word formation is the productive use of the Turkish suffixes *-li*, *-lik*, *-siz*, and *-çi* (cf. Masliyah 1996).

²⁷ For an Aramaic influence on syntax see above, section 4.3.

²⁸ For examples see Jastrow 1978: 36–38.

- (37) *?a-ğīd* *a-tšārək* *maʃā-k* *bə=hāda* *l-mūsəm*
 1SG-want.IPF 1SG-share.IPF with-2SG.M in=DEM.M.SG DEF-season
bə=z-zagəf
 in=DEF-crop
 “This season I want to share the crop with you.”
- (38) *qal=l-u* *l-maʃlāwi* *?ahlan* *w=sahlan* *bī-k*
 ‘say.PF.3SG.M=to-3SG.M DEF-Maşlāwi welcome in-2SG.M
 ‘The Maşlāwi said, “Welcome!”’
- (39) *wə=ttafaq* *maʃa* *š-šētān* *əz-zagəf*
 and=agree.PF.3SG.M with DEF-devil DEF-crop
yə-t.wazza' *bēn-āt-əm*
 3SG.M-PASS.devide.IPF between-PL-3PL
 ‘And he came to an agreement with the devil that the planted crop will be divided among them (as follows):’
- (40) *əl-fōq-āni* *mən* *əz-zagəf* *lə=l-maʃlāwi*
 DEF-above-NMLZ from DEF-crop to=DEF-Maşlāwi
wə=t-təħt-ēni *lə=š-šētān*
 and=DEF-under-NMLZ to=DEF-devil
 ‘The upper part of the planted crop for the Maşlāwi and the lower part for the devil.’
- (41) *w=wāfaq* *əš-šētān* *w=ğāh* *zagaʃ*
 and=agree.PF.3SG.M DEF-devil and=go.PF.3SG.M plant.PF.3SG.M
əl-maʃlāwi *ħənta*
 DEF-Maşlāwi wheat
 ‘The devil agreed and the Maşlāwi went to plant wheat.’
- (42) *w=ləmmən* *ṣāg* *əl-haʃād* *ğā*
 and=when become.PF.3SG.M DEF-harvest come.PF.3SG.M
š-šētān *də=y-āxəd* *ħəʃʃət-u*
 DEF-devil PURP=3SG.M-take.IPF share-3SG.M
 ‘When it was (the time) of the harvest, the devil came to take his share.’
- (43) *ğəšəf* *əl-maʃlāwi* *kən=axad* *əs-sənbəl*
 see.PF.3SG.M DEF-Maşlāwi AUX=take.PF.3SG.M DEF-ears
w=xallā=l-u *t-təbən*
 and=leave.PF.3SG.M=to-3SG.M DEF-straw
 ‘He saw that the Maşlāwi had taken the ears and had left him the straw.’

- (44) *əš-šētān axad=l-u şafni*
 DEF-devil take.PF.3SG.M=to-3SG.M meditation
w=qal=lu lɔ=l-maşlāwi
 and=say.PF.3SG.M=TO-3SG.M to=DEF-Maşlāwi
 ‘The devil pondered and (then) said to the Maşlāwi,’
- (45) *ət-ğīd nə-t.şārək marrt=əl-ləx ɻa=s-sani*
 2SG.M-want.IPF 1PL-RECP.share time=DEF-other on=DEF-year
lə-ğdīd-i
 DEF-new-F.SG
 ‘“Do you want us to share another time the next year?”’
- (46) *wə-l-maşlāwi ɻadd ɻalē-nu*
 DEF-Maşlāwi come.back.PF.3SG.M on-3SG.M
w=qal=l-u ɻē mā=yə-nşət ɻana
 and=say.PF.3SG.M=to-3SG.M yes NEG=3SG.M-ask.IPF 1SG
ħādəg
 to.be.ready.PTCP
 ‘The Maşlāwi responded to him, “Yes, why not? I’m ready!”’
- (47) *w=əš-šētān ɻəməl nafs-u šātəg*
 and=DEF-devil make.PF.3SG.M REFL-3SG.M clever
ɻa=l-maşlāwi w=qal=l-u
 on=DEF-Maşlāwi and=say.PF.3SG.M=to-3SG.M
 ‘The devil fancied himself to be more clever than the Maşlāwi and said to him,’
- (48) *ba.lākin ha-l-marra ɻana ɻa-ğīd āxəd*
 but DEM-DEF-time 1SG 1SG-want.IPF 1SG.take.IPF
əl-fōq-āni mən əz-zagħəf w=lə-k əğ-ğaww-ēni
 DEF-above-NMLZ from DEF-crop and=to-2SG.M DEF-under-NMLZ
 ‘“But this time I want the upper part of the planted crop and you will get the lower part.”’
- (49) *qəbel əl-maşlāwi w=ğāħ əš-šētān*
 accept.PF.3SG.M DEF-Maşlāwi and=go.PF.3SG.M DEF-devil
 ‘The Maşlāwi accepted and the devil went away.’
- (50) *w=ğā baħəd sani də=y-āxəd*
 and=come.PF.3SG.M after year PURP=3SG.M-take.IPF
ħəssət-u mən əz-zagħəf
 share-3SG.M from DEF-crop
 ‘After a year (the devil) came to take his share of the crop.’

- (51) *gəšəf* *əl-maslāwi* *kən=zagāf* *başal*
 see.PF.3SG.M DEF-Maşlāwi AUX=plant.PF.3SG.M onions
 ‘He saw that the Maşlāwi had planted onions.’
- (52) *əl-maşlāwi* *?axad* *həşşət-u* *ət-taht-ēnī-yi*
 DEF-Maşlāwi take.PF.3SG.M share-3SG.M DEF-under-NMLZ-F.SG
əl-başal *w=baqa* *lə=š-šētān* *bass* *əl-əʃğūš*
 DEF-onions and=remain.PF.3SG.M to=DEF-devil only DEF-leaves.PL
 ‘The Maşlāwi took his share, the lower part, the onions, and for the devil
 only the leaves remained.’

Abbreviations

AP	Active participle	MOD	Modifier
COLL	Collective	NEG	Negation
CONSTR.ST	Construct state	NMLZ	Nominalizer
COP	Copula	PASS	Passive
DEF	Definite article	PF	Perfective
DEM	Demonstrative	PL	Plural
DUAL	Dual	PRCL	Particle
ELA	Elativite	PTCP	Participle
F	Feminine	PURP	Purposive
GEN.MARKER	Genitive marker	REFL	Reflexive
IPF	Imperfectiv	SG	Singular
M	Masculine		

References

- Abu-Haidar, Farida. 2002. Negation in Iraqi Arabic. In Werner Arnold & Hartmut Bobzin (eds.), *Sprich doch mit deinen Knechten aramäisch, wir verstehen es!*. 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag, 1–13. Wiesbaden: Harrassowitz.
- Abu Haidar, Farida. 2004. The Arabic of Rabī'a: A *qəltu* dialect of north-western Iraq. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 1–12. Leiden & Boston: Brill.
- Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge: Center for Middle Eastern Studies of Harvard University.
- Contini, Riccardo. 1999. Le substrat araméen en néo-arabe libanais : Préliminaires à une enquête systématique. In Marcello Lamberti & Livia Tonelli (eds.), *Afroasiatica Tergestina: Papers from the 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics, Trieste, 23–24 April, 1998*, 101–122. Padova: Unipress.
- Holes, Clive. 2007. Colloquial Iraqi Arabic. In John Nicholas Postgate (ed.), *Languages of Iraq, ancient and modern*, 123–134. British School of Archaeology in Iraq.

- Jastrow, Otto. 1978. *Die mesopotamisch-arabischen qəltu-Dialekte, vol. 1, Phonologie und Morphologie*. Wiesbaden: Steiner.
- Jastrow, Otto. 1979. Zur arabischen Mundart von Mossul. *Zeitschrift für Arabische Linguistik* 2. 36–75.
- Jastrow, Otto. 1990. *Der arabische Dialekt der Juden von 'Aqra und Arbīl*. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 2004. Jüdisches, christliches und muslimisches Arabisch in Mossul. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 135–150. Leiden & Boston: Brill.
- Khan, Geoffrey. 1984. Object markers and agreement pronouns in Semitic languages. *Bulletin of the School of Oriental and African Studies* 47. 468–500.
- Khan, Geoffrey. 1997. The Arabic Dialect of the Karaite Jews of Hīt. *Zeitschrift für arabische Linguistik* 34. 53–102.
- Masliyah, Sadok. 1996. Four Turkish suffixes in Iraqi Arabic: -li, lik, siz and -çi. *Journal of Semitic Studies* 41. 291–300.
- Morony, Michael G. 2005. *Iraq after the Muslim Conquest*, 2nd ed. Piscataway, NJ: Gorgias Press.
- Procházka, Stephan. 2006–2007. Does geographical periphery imply linguistic periphery? The examples of the Arabic dialects of Cilicia and Urfa in southern Turkey. *Roman-Arabica* 6–7. 109–132.
- Socin, Albert. 1882–1883. Der arabische Dialekt von Mossul und Mardin. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*. 36: 1–53, 238–277; 37: 293–318.
- Talay, Shabo. 1999. *Der arabische Dialekt der Khawētna*. 1. *Grammatik*. Wiesbaden: Harrassowitz.
- Weninger, Stefan. 2011. Aramaic-Arabic language contact. In Stefan Weninger (ed.), *The Semitic languages: An international handbook*, 747–755. Berlin & Boston: De Gruyter Mouton.

Additional references

- Abu-Haidar, Farida. 1989. Maintenance and shift in the Christian Arabic of Baghdad. *Zeitschrift für Arabische Linguistik* 21. 47–62.
- Eksell, Kerstin. 2006. Analytic Genitive. In Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics*, vol. 1, 82–85. Leiden & Boston: Brill.
- Holes, Clive. 1995. Community, dialect and urbanization in the Arabic-speaking Middle East. *Bulletin of the School of Oriental and African Studies* 58. 270–287.
- Jankowski, Henryk. 1989. Word order in Iraqi Arabic. *Lingua Posnaniensis* 32–33. 109–112.
- Jastrow, Otto. 1981. *Die mesopotamisch-arabischen qəltu-Dialekte, vol. 2, Volkskundliche Texte in elf Dialekten*. Wiesbaden: Steiner.
- Jastrow, Otto. 1983. Tikrit Arabic verb morphology in a comparative perspective. *Al-Abhath* 31. 99–110.
- Jastrow, Otto. 1988. Zum arabischen Dialekt der Juden von Arbīl. *Zeitschrift für Arabische Linguistik* 19. 50–59.
- Jastrow, Otto. 1989. Notes on Jewish Maṣlāwi. *Jerusalem Studies in Arabic and Islam* 12. 282–293.

- Jastrow, Otto. 1990. Die arabischen Dialekte der irakischen Juden. In Werner Diem & Abdoldjavad Falaturi (eds.), *XXIV. Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln. Ausgewählte Vorträge*, 199–205. Stuttgart: Steiner.
- Jastrow, Otto. 1991. Brotbacken: Ein Text im arabischen Dialekt der Juden von Mossul in memoriam Ezra Laniado. *Zeitschrift für Arabische Linguistik* 23. 7–13.
- Jastrow, Otto. 1992. Der arabische Dialekt der Juden von Kirkuk. *Jerusalem Studies in Arabic and Islam* 15. 240–254.
- Jastrow, Otto. 1993. Geschichten aus Səndōṛ. *Zeitschrift Für Arabische Linguistik* 25. 161–177.
- Jastrow, Otto. 1994. The qəltu dialects of Mesopotamian Arabic. In Jorge Aguadé, Federico Corriente & Marina Marugán (eds.), *Actas del congreso internacional sobre interacciones lingüísticas arabo-romances y paralelos extra-iberos*, 119–123. Zaragoza.
- Jastrow, Otto. 2001. Three anecdotes in Jewish Maslawi. In Judith Rosenhouse & Ami Elad-Bouskila (eds.), *Linguistic and cultural studies on Arabic and Hebrew. Essays presented to Professor Moshe Piamenta for his eightieth birthday*, 61–70. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 2005. Linsen mit Reis – und doch keine mžaddara! Jüdisch-arabische Kochrezepte aus Mossul. In Thomas Bauer & Ulrike Stehli-Werbeck. (eds.), *Alltagsleben und materielle Kultur in der arabischen Sprache und Literatur: Festschrift für Heinz Grotfeld zum 70. Geburtstag*, 223–232. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 2006. Where do we stand in the research on the Anatolian qəltu dialects? *Roman-Arabica* 6–7. 63–69.
- Jastrow, Otto. 2007. Iraq. In Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics*, vol. 2, 414–424. Leiden & Boston: Brill.
- Krotkoff, Georg. 1991. Remarks on the Iraqi-Arabic lexicon: problems and characteristics. In Alan S. Kaye (ed.), *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*, 886–889. Wiesbaden: Harrassowitz.
- Lecerf, Jean. 1967. Regards sur l'état présent des études sur les dialectes arabes de Bagdad. In *Atti del terzo Congresso di studi arabi e islamici, Ravello 1966*, 443–459. Naples: Istituto Universitario Orientale.
- Levin, Aryeh. 1994. The role of the Mesopotamian qəltu dialects in the history of the Arabic dialects of Iraq. In Dominique Caubet & Martine Vanhove (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris. Colloque international tenu à Paris du 27 au 30 janvier 1993*, 317–332. Paris: Publications Langues’O.
- Levin, Aryeh. 2000. The ‘imāla in the modern qəltu dialects. In Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi & Zakia Iraqui-Sinaceur (eds.), *Aspects of the dialects of Arabic today. Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA), Marrakech, Apr. 1–4, 2000*, 98–104. Rabat: Amapatril.
- Mansour, Jacob. 1985. The particle *baqa* in the Judaeo-Arabic dialect of Baghdad. *Zeitschrift für Arabische Linguistik* 14. 62–75.
- Mansour, Jacob. 2001. The particles ‘aš/‘eš and ‘ášu in the Jewish Arabic of Baghdad. *Zeitschrift für Arabische Linguistik* 39. 84–89.
- Masliyah, Sadok. 1997. The diminutive in spoken Iraqi Arabic. *Zeitschrift für Arabische Linguistik* 33. 68–88.
- Melcer, Ioram. 1995. The analytical genitive in the Jewish-Baghdadi Arabic dialect. *Zeitschrift für Arabische Linguistik* 29. 59–76.

- Müller-Kessler, Christa. 2003. Aramaic ‘*k*’, *lyk*’ and Iraqi Arabic ‘*aku*, *māku*: the Mesopotamian particles of existence. *Journal of the American Oriental Society* 123. 641–646.
- Retsö, Jan. 1987. Copula and double pronominal objects in some Semitic Languages. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 137. 229–236.
- Retsö, Jan. 2004. Relative-clause marking in Arabic dialects: a preliminary survey. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects: a collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 263–273. Leiden & Boston: Brill.
- Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization*. Winona Lake: Eisenbrauns.
- Salonen, Erkki. 1980. *On the Arabic dialect spoken in Širqāt (Assur)*. Helsinki: Suomalais-Ugrilaisen Seuran julkaisuja 100. Suomalais-Ugrilaisen Seuran toimituksia 100.
- Zemer, Hila. 2008. The copula in the Christian Arabic dialect of Baghdad – possible origins. In Stephan Procházka & Veronika Ritt-Benmimoun (eds.), *Between the Atlantic and Indian Oceans: Studies on contemporary Arabic dialects. Proceedings of the 7th AIDA Conference*, 505–514. Vienna & Münster: LIT.

3.3. The Iranian languages of northern Iraq

Geoffrey Haig

1. Introduction

The Iranian languages currently spoken in northern Iraq can be assigned to three main groups: Behdinī, the local name for the varieties of Northern Kurdish (or Kurmanjī) spoken in Iraq; Sorani Kurdish (used here synonymously with Central Kurdish), and Gorani.¹ The latter subsumes Hawrāmī of the Halabja region, together with a cluster of other dialects that go under various names. All three groups are traditionally classified as belonging to the northwestern branch of Iranian. The approximate locations of Northern and Central Kurdish are provided in Fig. 1, while Gorani varieties are shown in Fig. 4 below.

Both Behdinī and Sorani are unanimously considered varieties of Kurdish, while the nature of the relationship between Kurdish and Gorani remains a matter of controversy. As will become apparent, at least in terms of morphology, the various varieties of Gorani diverge from Sorani and Behdinī rather systematically, implying a historically more distant relationship. From the perspective of comparative Iranian philology, then, a distinction is justified. Culturally and in terms of subjective identity perceptions among the speech communities, on the other hand, there are reasons for including Gorani within a broader socio-cultural notion of “Kurdish” (see Haig and Öpengin 2014 on the concept of “Kurdish”).

Among the three groups, the most important in terms of prestige, degree of standardization, media representation, and number of speakers, is Sorani (MacKenzie 1961, 1962), spoken by around three million speakers in Iraq, with a further three million in neighbouring regions of Iran, and further north to the shores of Lake Urmi (also spelled Urmiye). The least-well documented languages are the various varieties of Gorani. In particular, very little reliable information is available on the fragmented groups of dialects spoken westward of the Hawrāmī region, variously referred to as Kaka’ī, Šabak, Sarlī, or Bājałānī (see §5). The majority of Sorani and Behdinī speakers are Sunnite Muslims, with the exception of the Ezidi communities among the Behdinī speakers, while Gorani speakers are generally affiliated with heterodox, or non-Islamic, religious beliefs.

¹ There is little agreement with regard to the spelling conventions for language names, and this overview makes no claims to consistency. Generally the form that is most widely-used in English publications has been chosen, avoiding diacritics and non-standard characters, but in the case of lesser-known varieties, a form more closely representing a transcription is chosen.

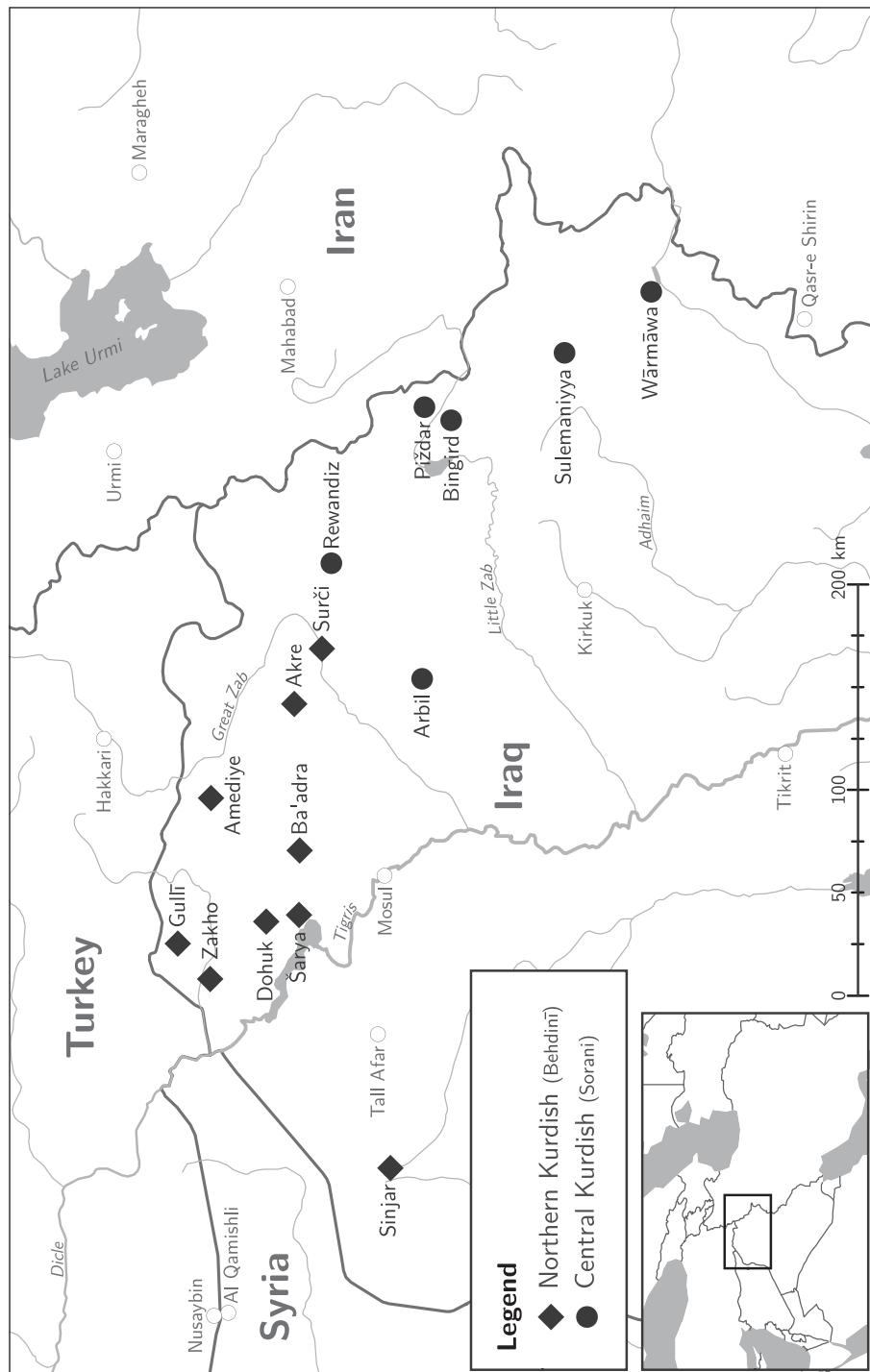


Figure 1: Approximate locations of Northern and Central Kurdish varieties mentioned in this chapter

This chapter gives a brief overview of the main structural features of each, together with some comparative notes. Regrettably, since the pioneering work of MacKenzie on Kurdish dialects of Iraq, undertaken in the 1950's, and Blau's study of Behdinī (1975), very little research on the vernaculars has been undertaken, and the coverage here is correspondingly uneven.

2. Background to the speech communities

When compared to Neo-Aramaic, the degree of early linguistic attestation of the Iranian languages of Iraq is meagre. For Kurdish generally, textual attestation does not extend farther back than around the 16th century. Nor can the predecessor of Kurdish be equated with any of the historically attested Western Middle Iranian languages. Standard accounts of the history of the Kurds (McDowall 2004) locate Kurdish tribes in Mesopotamia from the outset of the Islamic expansion in the seventh and eighth centuries, prior to the establishment of larger Kurdish principalities in the Jazira in the 10th and 11th centuries (McDowall 2004: 23; see also Haig, this volume, chapter 2.3, §1.1). Up until the 19th century, the destiny of the Kurdish groups in the region was shaped by the shifting fortunes of Ottoman/Iranian rivalry, with the majority of Kurdish tribes supporting the former, while never entirely abandoning their independence. Two major groupings can be identified: the Ardalan principality, and the Baban Kurds. The Ardalan controlled large amounts of territory both eastward and westward of the Zagros mountains in the fourteenth and fifteenth centuries, with their centre located at Senna (today's Iranian city Sanandaj). Their allegiance lay initially with the Iranian rulers, though they shifted to the Ottomans over the 16th and 17th centuries, and thus retained some influence in Ottoman-dominated regions of today's northern Iraq. The Ardalans are associated with the Gorani language, and with the Ahl-i Haqq (or Yaresan) religion, which accounts for the presence of groups affiliated with these beliefs in today's northern Iraq (see §5). There is a considerable corpus of poetry in the written koiné form of Gorani, dating back to the 14th century (MacKenzie 2002), while written prose is almost completely lacking. The other major tribal power, rivals to the Ardalan, was the Baban (*Bābān*) confederacy, centred around Sulaymaniyah, which became a centre of Kurdish cultural and linguistic development. Primarily allied with the Ottomans, they maintained a dominant role in today's northern Iraq up until the 19th century, when the Babans were largely disbanded by the Ottomans during their efforts at reformation. This effectively put an end to the Kurdish emirates as major political and military actors in the region (McDowall 2004: 47). The language of the Babans was primarily Sorani.

Among the most fascinating aspects of northern Iraq's linguistic landscape is the admixture of heterodox Islamic, and non-islamic groups concentrated in the region north and east of Mosul. Along with the Christian and Jewish NENA

communities (see Khan, this volume, chapter 3.4), the region also hosts Northern Kurdish-speaking Ezidī communities as well as a range of Yaresan-affiliated religious groups, and others exhibiting what Leezenberg (1994) refers to as “pre- and non-islamic belief elements”. However, their histories remain shrouded in mystery, as they “developed outside the major centers of the Islamic world” (Leezenberg 1994), and tend, through necessity, to maintain a low profile. The ravages of the Islamic State, which effectively occupied the city of Mosul from 2014–2017, but which had been persecuting non-orthodox religious groups for several years previously, has now largely obliterated much of the historical diversity of this region.

Sorani and Behd. are both widely spoken in their respective regions, are widely used in broadcasting and internet media, are acquired by children as a first language, and are represented in the education system of the Autonomous Kurdish Region of Iraq, though to differing degrees. Of the two, Sorani has the highest prestige, and the best-established written tradition (which uses a modified version of the Arabic script). Behd., on the other hand, has a more troubled status: it is the less prestigious variety of Kurdish in Iraq, thus ranges behind both Sorani and Arabic in terms of overall status, nor does it profit from the emergent standardisation of Kurmanjî, which is based on a Roman-alphabet writing system and propagates a standard form that differs quite significantly from Behd. (see Haig and Mustafa, in press). The Gorani varieties, on the other hand, have virtually no official status, and must be considered highly endangered.

3. Sorani (Central Kurdish)

Sorani, also known as Central Kurdish (CK) refers to a range of dialects spoken in Iraq and the bordering regions of Iran (see Fig. 1 in Haig, this volume, chapter 2.3). It has been quite extensively documented and analysed in grammatical descriptions and dictionaries, and has a well-established written standard language, using a modified version of the Arabic script, based on the dialect of the northern Iraqi city of Sulemaniyya (see Hassanpour (1992: Ch.8) on the emergence of standardized Sorani). A structural overview of Sorani, based on the Sulemaniyya dialect, is provided in McCarus (2009), while Öpentin (2016) provides a comprehensive linguistic analysis of the dialect of Mukri, spoken near the city of Mahabad in West Iran. The most detailed treatment from the perspective of dialectal variation remains MacKenzie (1961 and 1962), covering several Central Kurdish dialects spoken in northern Iraq. This overview focusses on the spoken vernacular, rather than the standard language, and draws primarily on MacKenzie (1961 and 1962), McCarus (2009), and Öpentin (2016).

3.1. Phonology

3.1.1. Consonants

According to MacKenzie (1961: 1), the Sulemaniyya dialect has 28 consonant phonemes, shown in Table 1.

Table 1: Sulemaniyya Sorani consonant phonemes

	Labial	Alveo-dental	Post-alv.	Palatal	Velar	Uvular	Pharyn.	Glott.
Stop	p b	t d			k g	q		(?)
Affricate				ʃ dʒ				
Nasal	m	n			ŋ			
Fricative	f v	s z	ʃ ʒ		x ɣ		χ ħ	h
Tap/Trill			f / r					
Approximant								
Lateral		l	t					
Glide	w		j					

Sorani lacks the phonemic three-way stop distinction typical of Northern Kurdish (voiceless aspirated / voiceless non-aspirated / voiced, see Haig, this volume, chapter 2.3). Pharyngealized segments outside of the pharyngeal fricatives are not included as phonemes here, although regularly realized in Arabic loans and in varying degrees in the inherited lexicon. An important feature of Sulemaniyya Kurdish is the so-called ‘Zagros d’. This refers to a general instability of /d/ when following a sonorant, except /r/. The result may be the lenition of /d/ to a vocalic segment, or assimilation to a preceding sound, or loss: *birñdar* [bri:nna:r] ‘wounded’; *bad* [baə] ‘bad’. In other environments, /d/ is preserved: *dast* [dast] ‘hand’, *kird* [kird] ‘did’. It is noteworthy that the stem initial /d/ of one of the most frequent verbs, *dān* ‘give’, is regularly lenited, even when not preceded by a sonorant. Presumably this is due to the fact that most inflected forms of this verb involve a sonorant preceding the stem (e. g. *a-y-da-m* IND-3s.P-give.prs-1s [ɛi.jem] ‘I give it’ (MacKenzie 1961: 3), and the resulting lenition has generalized to all forms of this verb. Dialectal variation in phonology affects mostly allophonic realization rather than the underlying phonemic system. Of note is the realization of /k, g/ as the affricates [ʃ, dʒ] respectively, and the fronting of the original affricates to [tʃ, dz] (alveo-palatal affricates), notable in Bingird and Piždar dialects (MacKenzie 1961: 24–25). In fact palatalization of /k, g/ before front vowels is widespread across Sorani, and yields affricates in many of the dialects north of Sulemaniyya (e. g. in Mukri, Öpengin 2016).

3.1.2. Vowels

According to McCarus (2009: 591), Sulemaniyya Central Kurdish has the nine vowel phonemes provided in Fig. 2.

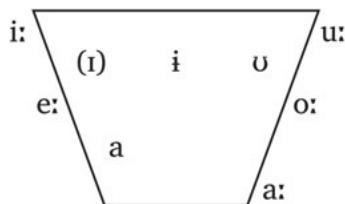


Figure 2: Vowels of Sorani

MacKenzie (1961: 1) gives a slightly different account of the vowels, which includes a mid front rounded [ø], and only one short unrounded central vowel. The first could be analysed as [w] plus a front vowel (as in McCarus 2009), but see Öpengin (2016) for a different view. The central unrounded [ɪ] in Fig. 2, not included in MacKenzie (1961), is mainly found in Arabic loans such as [imtihān] ‘examination’ (McCarus 2009: 591), and is thus placed in brackets here. The vowel shown here as [a] has a broad range of realizations, most frequently as schwa (before glides within the same syllable), [æ], [ɛ] before [j] in the following syllable [he.jə] ‘there is’, and low central [a] adjacent to pharyngealized [š], as in [šaṣt] ‘sixty’ (McCarus 2009: 591). In keeping with the conventions used below for Behdinī and those of Öpengin (2016) for the Mukri dialect of Central Kurdish, but unlike MacKenzie (1961) and McCarus (2009), I transcribe it here with <e>.

Syllable onsets are relatively complex for a West Iranian language; one might conjecture Semitic influence here. According to McCarus (2009: 593), any combination of two consonants is permitted in syllable onsets, though certain restrictions obtain (e. g. stops are only licensed as second consonants when the first is also a stop): *ktēb* ‘book’, *tfej* ‘rifle’, *xrāp* ‘bad’. However, these complex onsets can generally be analysed as resulting from elision of underlying [i], still audible in careful speech. Nevertheless, it is striking that Sorani speakers evidently have no difficulty articulating such complex onsets, which are undoubtedly the norm in natural speech. CCC-onsets are not reported (the cases of CCC-clusters in McCarus (2009: 593) either involve an intervening syllable boundary, or an initial glide). Syllable-internal vowel combinations are avoided within individual lexical items, but may arise through combinations at phrase level.

3.2. Morphosyntax

Although Central Kurdish is regularly considered a ‘dialect’ of ‘Kurdish’, the differences between Central and Northern Kurdish in the realm of morphology are considerable, and heavily impede mutual intelligibility (Öpentin and Haig 2014; Haig and Öpentin 2018). In contrast to Northern Kurdish, most of Central Kurdish (i) lacks grammatical gender on nouns; (ii) lacks morphological exponence of structural case; (iii) has a suffixed marker of definiteness *-eke*; (iv) has a distinct paradigm of clitic personal pronouns, not derivable from the free pronouns; (v) has an affixal passive marker. In the Sulemaniyya dialect, the passive marker attaches to the present stem of the verb, and then secondarily differentiates tense: *kuštin* ‘kill’ (infinitive) > *kuž-* (present stem) > *kuž-rē-* (passive, present), *kuž-rā-* (passive, past). The realization of passive morphology varies considerably across Central Kurdish (MacKenzie 1961: 118–119).

From differences (i) and (ii), it follows that the Ezafe particle is undifferentiated for gender; it is also undifferentiated for number. Structural case relations (S, A, P, and possessive) are carried via word-order (basically SV, AOV, N-possessor within the NP), rules of agreement morphology that are sensitive to grammatical relations, and the clitic pronouns. Other case relations (local, non-core arguments such as instruments or comitatives etc.) are expressed through a rich inventory of pre- and circumpositions. As in most of West Iranian, verbs constitute a small lexical class, which is primarily expanded through complex predicate-formation (non-verbal element plus a light verb, for example *kirdin* ‘do’, *dān* ‘give’, *būn* ‘be, become’, *birin* ‘take’, *kewtin* ‘fall’, *hātin* ‘come’). The non-verbal element may be transparently related to a lexical item that occurs in other contexts (e. g. *teslīm kirdin* ‘surrender, hand over’, or it may be a particle of uncertain word-class membership, as in *hāl wāsīn* ‘hang up’, with the particle *hāl* indicating approximately ‘upward’).

3.2.1. Nominal morphology

From a broader Iranian perspective, one of the most striking features of Sorani nominal morphology is the suffixal expression of definiteness. While indefiniteness suffixes are widespread across west Iranian, only a small pocket of Iranian languages spoken in northern Iraq and neighbouring regions of Iran have developed definiteness suffixes in addition to the indefiniteness markers. Table 2 (adapted from McCarus 2009: 598) gives an overview of the morphological marking of definiteness.

Table 2: Indefiniteness and definiteness in Sorani

	GENERIC, UNDER-SPECIFIED FOR NUMBER	INDEFINITE	DEFINITE	PROXAL/DISTAL DEMONSTRATIVE
SG	<i>pyāw</i> ‘man, men’	<i>pyāw-ēk</i> ‘a man’	<i>pyāw-eke</i> ‘the man’	<i>em / ew pyāw-e</i> ‘this/that man’
PL		<i>pyāw-ān</i> ‘men’	<i>pyāw-ek-ān</i> ‘the men’	<i>em / ew pyāw-ān-e</i> ‘these/those men’

This definiteness suffix, with reflexes across Gorani and in Southern Kurdish, is of considerable interest from a diachronic perspective. Cross-linguistically, definiteness markers (often termed articles) are often the result of grammaticalization of some earlier independent deictic element (typically a demonstrative, see e. g. Skrzypek 2010). But for Sorani, no plausible candidate source element for such a grammaticalization process is available. Rather, it appears the suffix goes back to a derivational suffix, widely attested in the Western Middle Iranian languages Middle Persian and Parthian; see Jahani (2015) and Haig (in press) for discussion. It is nevertheless worth noting that the definiteness suffix is not equivalent to, for example, the definite article in English: it is not used with unique referents, like *the sun*. Furthermore, it is not even consistently present on nouns that have discourse-recoverable referents. The available original texts have numerous examples of non-suffixified nouns occurring in contexts where one would expect a definiteness suffix (for example ‘the letter’ in MacKenzie’s narrative text (1962: 10). I would therefore hesitate to refer to this marker as a ‘definite article’. This is an area requiring much more research.

In combination with demonstratives, NPs take an additional stressed -é. When demonstratives are used without complements, they too take the demonstrative marker (the clitic/suffix distinction is not addressed here), as shown in Table 3.

Table 3: Sorani pronouns, personal and demonstrative

	PERSONAL			DEMONSTRATIVE	
	1ST PERS.	2ND PERS.	3RD PERS.	‘that’	‘this’
SG	<i>min</i>	<i>tō</i>	<i>ew</i>	<i>ew-e</i>	<i>em-e</i>
PL	<i>ēme</i>	<i>ēwe</i>	<i>ew-ān</i>	<i>ew-ān-e</i>	<i>em-ān-e</i>

As mentioned, Sorani also has an additional set of clitic personal pronouns. Section 3.2.4 below is dedicated to the morphosyntax of clitic pronouns.

3.2.1.1. Ezafe constructions

NPs are generally head-initial, with attributes following, but demonstratives and numerals precede the head noun. Post-nominal attributes are linked to the head by an Ezafe-clitic, of which two kinds can be distinguished:

The phrasal Ezafe =*ī* (EZ)

This is used to freely combine phrasal elements to a complex NP. The dependent (second element) can be a nominal phrase of any category (AP, NP, PP, or pronoun). The head may carry the indefinite suffix -*ēk* (cf. Table 3), i. e. the presence of an attribute does not preclude indefiniteness of the head. The phrasal Ezafe can be repeated when the head is accompanied by more than one attribute, as in (4) below:

- (1) *le xošī=y ew-e*
from pleasure=EZ that-DEM
'from the pleasure of that' (MacKenzie 1961: 62)
- (2) *tūtik-ēk=ī pičkole*
dog-INDF=EZ small
'a little dog' (MacKenzie 1961: 63)
- (3) *xānū=y ēme*
house=EZ 1PL
'our house' (MacKenzie 1961: 63)
- (4) *kič-ēk=ī džwān-ī čwārde sāl*
girl-INDF=EZ beautiful=EZ fourteen year
'a beautiful girl of fourteen years' (MacKenzie 1961: 63)

The compound Ezafe =*e* (CEZ)

This element is used in frequent and partially lexicalized combinations, but also for freer and evidently not lexicalized combinations. I nevertheless refer to it as a compound ezafe, following McCarus (2009: 613), for mainly formal reasons: First, the entire construction may be marked with the definiteness or indefiniteness suffixes -*eke*, or -*ēk*, while in none of the examples I have encountered is the head noun itself inflected for definiteness or indefiniteness. There are thus restrictions on internal inflection, which are strongly suggestive of compound status. However, it must be noted that many of the examples cited in MacKenzie (1961: 64) are semantically incongruous for compounds; more research is required on this topic.

- (5) *kič=e jwān-eke*
girl=CEZ pretty-DEF
'the pretty girl' (McCarus 2009: 613)

- (6) *kilk=e ker-ēk*
tail=CEZ donkey-INDF
‘a donkey’s tail’ (MacKenzie 1961: 65)
- (7) *ew xēwet=e sewz-e*
that tent=CEZ green-DEM
‘that green tent’ (MacKenzie 1961: 64)

3.2.2. Verbal morphology

As in other West Iranian languages, each verb has two stems, called here the past and present stem respectively. Examples of stems for common verbs are provided in Table 4. The infinitive, the traditional citation form for Kurdish verbs, is formed from the past stem via the addition of *-(i)n*. In natural speech, infinitives are only rarely attested, barring certain lexicalized forms.

Table 4: Past and present stems of frequent verbs in Sorani

GLOSS	INFINITIVE	PAST STEM	PRESENT STEM
go	<i>čūn / rōištin</i>	<i>čū- / rō-</i>	<i>č-</i>
come	<i>hātin</i>	<i>hāt-</i>	<i>(h)ē-</i>
be	<i>būn</i>	<i>bū</i>	<i>b-</i>
give	<i>dān</i>	<i>dā-</i>	<i>de-</i>
do, make	<i>kirdin</i>	<i>kird-</i>	<i>ke-</i>
see	<i>dīn</i>	<i>dī-</i>	<i>bīn-</i>
fall	<i>keftin, kewtin</i>	<i>kewt, keft-</i>	<i>kew-</i>
say	<i>wutin</i>	<i>wut-</i>	<i>tē-</i>
eat	<i>xwārdin</i>	<i>xwārd-</i>	<i>xō-</i>

Each stem forms the base for a particular set of TAM and negation affixes, constituting a number of distinct paradigms. Very roughly, those based on the past stem are associated with past time reference, and to some extent with past unrealis modality, while present stems form indicative present, future, and various kinds of unrealized (or non-asserted) forms, typically in complement clauses governed by verbs such as ‘hope’, ‘want’, ‘intend’ etc.

Predicates are obligatorily indexed for person and number of one argument. Three distinct paradigms can be identified, shown in Table 5. Generally, the indexed argument is the S or A, but with transitive verbs in the past tense, verbal affixes may index the P (if it is not otherwise expressed in the clause), or under specific conditions, an indirect participant, cf. §3.2.4, ex. (18), or may be simply

default third person singular if the P is overtly present in the clause; at this point, details of person marking vary considerably across and even within dialects, see Öpelin (2016) for detailed discussion. The copular clitics attach to non-verbal predicates in the present tenses, but are also used with certain verb forms, most notably participles (see Table 7 below).

Table 5: Verbal affix and copular person markers in Sulemaniyya Sorani

VERBAL AFFIXES				COPULAR CLITICS
	PRESENT STEM	PAST STEM		
SG	1 -im	-im	=im	
	2 -ī(t), -y(t)	-ī(t)	=ī	
	3 -ē(t) ²	-Ø	=e	
PL	1 -īn	-īn	=īn	
	2 -in	-in	=in	
	3 -in	-in	=in	

3.2.2.1. Verb forms based on the present stem

Finite verb forms based on the present stem of the verb require one of the six mutually exclusive present-tense TAM and negation prefixes, though in imperatives, the prefix may be dropped. The appropriate form of the present stem person markers (cf. Table 5) is suffixed to the stem. The prefixes are shown in Table 6.

Table 6: TAM and negation prefixes with the present stem

TAM/Negation (abbreviation)		Example with present stem of <i>wutin</i> ‘say’
indicative (IND)	e-	<i>e-łē-m</i> ‘I say, am saying’
negated (NEG)	nā-	<i>nā-łē-m</i> ‘I don’t say’
subjunctive (SUBJ)	bi-	<i>bi-łē-m</i> ‘that I say’
subjunctive negated (SUBJ.NEG)	ne-	<i>ne-łē-m</i> ‘that I not say’
imperative (IMP)	bi-	<i>bi-łē</i> (2SG), <i>bi-łē-n</i> (2PL) with consonant-final stem (e. g. ‘go’): <i>bi-č-e</i> (2SG), <i>bi-č-in</i> (2PL)
imperative negated (IMP.NEG)	me-	<i>me-łē</i> , <i>me-łē-n</i>

² According to MacKenzie (1961: 90), the “euphonic” -t of the second and third person singular is realized before vowel-initial clitics, or before a pause, but otherwise left unrealized.

The form of the present indicative is an important isogloss within Central Kurdish, with Sulemaniyya and Wārmāwa dialects having *e-*, but elsewhere *de-* (the former is adopted for the standardized written language). The forms of the other TAM/negation prefixes is fairly consistent across all dialects, and have obvious cognates throughout most of West Iranian. All TAM/negation prefixes are potential hosts for pronominal clitics, and may thus be separated from the verb stem:³

- (8) *e=mān=ewē* *bi-řō-yn*
 IND=1PL=be.desirable.PRS.3SG SUBJ-go.PRS-1PL
 ‘We want to go’ (McCarus 2009: 620)

- (9) *bi=y=gir-in*
 SUBJ=3SG=hold.PRS-2PL
 ‘Hold it!’ (MacKenzie 1961: 93)

Indicative forms do not distinguish progressive/immediate from habitual senses. Unlike Northern Kurdish, Sorani lacks a dedicated future marker, with the present tense used in future contexts, when a reasonable degree of certainty of fulfillment is implied. The subjunctive verb forms are used for dependent clauses, particularly following predicates of desire (as in (8) above), ability, and obligation, and more generally in independent clauses to express events that are not asserted, but are portrayed as possible, hypothetical, or desired. These may be introduced by modal particles such as *bā* ‘let’s’, or an expression such as *hāyfe* ‘it’s a pity’:

- (10) *bā mināl=ī tō bi-bīn-im*
 MOD child=EZ 2SG SUBJ-see.PRS-1SG
 ‘Let me see your children’ (MacKenzie 1961: 106)

- (11) *hāyf=e bi=y=kuz-īn*
 pity=COP.3SG SUBJ=3SG=kill.PRS-1PL
 ‘It’s a pity that we should kill it’ (MacKenzie 1961: 77)

3.2.2.2. Verb forms based on the past stem

Unlike the present stem, a number of verb forms based on the past stem are unprefixed, including the simple past. However, the past stem is also the base for a participle, in Sulemaniyya of the form *-uw* (on MacKenzie’s (1961: 97) analysis), e. g. *hāt-uw* ‘come.PST-PTCPL’, which in turn forms the base for a large number of secondary verb forms involving copular forms of the person markers (see Table 5).

³ The ability of TAM and negation prefixes (or perhaps proclitics) to host pronominal clitics is an important isogloss distinguishing Central Kurdish from Southern Kurdish and Gorani; in the latter, these prefixes are not possible hosts for pronominal clitics.

The analysis of McCarus (2009) and MacKenzie (1961) differ on a number of points, which cannot be disentangled here. I will only present a selection of those verb forms, following MacKenzie (1961), which appear to be uncontroversial, while referring to the sources for more details. Table 7 provides an overview of the six paradigms introduced below.

I Indicative forms based on the past stem

Ia Simple past

Consists of the past stem plus the past-stem person markers from Table 6.

Ib Past Imperfective

Identical to the simple past, but with the addition of the prefix *e-* (with dialectal variants, as discussed above).

Ic Perfect

Based on the participle Past stem *-uw*, to which the enclitic forms of the copular person indexes are added.

Id Pluperfect

Based on the past stem, to which an inflected simple past form of the verb *būn* ‘be’ is attached. Following a stem-final consonant, an epenthetic vowel [i] intervenes between verb stem and the form of *būn* (this epenthetic vowel is lacking in dialects other than Sulemaniyya).

II Subjunctive and conditional forms based on the past stem

IIa The past conditional (MacKenzie 1961: 97)

Based on the simple past form of the verb, to which the subjunctive prefix *bi-* is added, and the suffix (clitic?) *-āye* following the person markers.

IIb Perfect conditional (MacKenzie 1961: 99–100)

Identical to the pluperfect, except that the form of the verb *būn* ‘be’ which is attached to the verb is the present subjunctive, not the simple past.

Table 7: TAM forms based on the past stem, illustrated with the verb *hātīn* ‘come’

	INDICATIVE				NON-INDICATIVE	
	SIMPLE PST.	IMPF.	PERF.	PLUPERF.	CONDITIONAL	PERF. SUBJ.
1SG	<i>hāt-im</i>	<i>e-hāt-im</i>	<i>hāt-uw-im</i>	<i>hāt-i-bū-m</i>	<i>bi-hāt-im-āye</i>	<i>hāt-i-bim</i>
2SG	<i>hāt-ī(t)</i>	<i>e-hāt-ī(t)</i>	<i>hāt-uw-ī(t)</i>	<i>hāt-i-bū-ī(t)</i>	<i>bi-hāt-ī(t)-āye</i>	<i>hāt-i-bī(t)</i>
3SG	<i>hāt</i>	<i>e-hāt</i>	<i>hāt-uw-a</i>	<i>hāt-i-bū</i>	<i>bi-hāt-āye</i>	<i>hāt-i-bē(t)</i>
1PL	<i>hāt-īn</i>	<i>e-hāt-īn</i>	<i>hāt-uw-īn</i>	<i>hāt-i-bū-īn</i>	<i>bi-hāt-īn-āye</i>	<i>hāt-i-bīn</i>
2/3PL	<i>hāt-in</i>	<i>e-hāt-in</i>	<i>hāt-uw-in</i>	<i>hāt-i-bū-in</i>	<i>bi-hāt-in-āye</i>	<i>hāt-i-bin</i>

The uses of the different TAM forms conform approximately to the expectations conveyed by the traditional labels, but see MacKenzie (1961: 133–140) for more details.

3.2.3. Word order in the simple clause

Word order in Sorani is largely identical to that of Northern Kurdish, both within the NP and the simple clause: SV, AOV. Like Behdinī, Sorani also makes extensive use of post-predicate goals, generally flagged through a clitic particle $=\{y\}e$, attached to the verb and glossed DRCT, for the goals of motion verbs. This particle is probably the historical reflex of an older preposition that flagged the goal NP itself, but it is now realized as a clitic to the verb:

- (12) *būk=yān hēnā=ye māl=ewe*
 bride=3PL bring.PST=DRCT home=ASP
 ‘They brought the bride back home’ (MacKenzie 1962: 62)

The directional particle is absent, however, when the post-predicate goal is governed by the prepositions *be* and *bō*:

- (13) *hal=tār girt be āsmānā*
 upwards=3SG take.PST into sky
 ‘She took him up into the sky’ (MacKenzie 1962: 46)
- (14) *kem=tān čū-w-in bō henār*
 which=2PL.POSS go.PST-PRF-2PL for pomegranates
 ‘Which of you went for the pomegranates?’ (MacKenzie 1962: 52)

Despite the placement of most goal arguments after the verb, the placement of direct objects is fairly consistently in pre-verbal position, so that Sorani can still be characterized as OV. Fronting of an object to a clause-initial position is possible, but post-posing an object after the verb is scarcely attested in the available texts.

3.2.4. Clitic personal pronouns

The clitic forms of the personal pronouns are a feature common to most of contemporary West Iranian, with a few exceptions such as Northern Kurdish, Zazaki, and Sangesar (Windfuhr 2009: 33). The forms are provided in Table 8.

Table 8: Clitic forms of personal pronouns in Sorani

	1ST PERS.	2ND PERS.	3RD PERS.
SG	$=\{i\}m$	$=\{i\}t$	$=\bar{t} / =y$
PL	$=mān$	$=tān$	$=yān$

Within the languages conventionally assigned to ‘Kurdish’ (Haig and Öppling 2014), the presence versus absence of clitic pronouns is a major isogloss, with Sorani, Gorani, and Southern Kurdish exhibiting them, while Kurmanjî and Zazaki lack them. There is a narrow belt of overlap, e. g. the dialect of Surči in North Iraq (MacKenzie 1961: 222), but on the whole, Kurdish seems either to make abundant use of the clitic pronouns, or abandon them entirely. In Sorani, they are used in the following five functions, labeled A-E.

A

Pronominal possessor, where they attach to the possessed NP:

birā-k-ān=im brother-DEF-PL-1SG ‘my brothers’ (MacKenzie 1962: 6)

B

Pronominal complement of adpositions:

leget=tān with=2PL ‘with you(PL)’ (MacKenzie 1962: 10)

C

Direct object of a verb in the present tense:

attaches to the right edge of the first available stress-bearing constituent of the VP (often a TAM or negation prefix on the verb, e. g. (11) above):

- (15) *be kuř-e pāšā=y nā-de-yn*
 to son-EZ king=3SG NEG-GIVE.PRS-1PL
 ‘We are not giving **her** to the Pasha’s son’ (MacKenzie 1962: 12)

D

The A of a past-tense transitive verb:

attaches to the first stress-bearing constituent of the VP. Three possible hosts, including the verb itself, are illustrated in (16a–c):

- (16) a. *min seg-eke=m ne-kušt*
 1SG dog-DEF=1SG NEG-kill.PST(3SG)
 ‘I didn’t kill the dog’
 b. *min ne=m=kušt*
 1SG NEG=1SG=kill.PST(3SG)
 ‘I didn’t kill (it)’
 c. *min kušt=im*
 1SG kill.PST(3SG)=1SG
 ‘I killed (it)’ (Haig 2017)

If a direct object is fronted to the left of a subject, e. g. for pragmatic purposes, it cannot host the A-clitic. Because the subject itself is by definition outside of the VP, in a clause where the object precedes the subject, neither can host the clitic, which would then typically go to the verb, as in the following:

- (17) *kuř-ēk=mān seg bird-uw-yat=ī*
 boy-INDF=1PL.POSS dog take.PST-PRF-3SG=3SG(A-CLITIC)
 ‘a boy of ours, the dog took (him)’ (MacKenzie 1962: 38)

E

The Indirect Participant (the human, non-affected, non-agentive, non-core argument of a transitive or intransitive clause, typically a recipient, a benefactive, a ‘wanter’, or an external possessor):

If not linked to an adposition, or blocked by another argument-indexing clitic (see §3.2.4.1), clitic placement is similar to the rule above (first stress-bearing constituent of the VP) but may include subjects:

- (18) a. *řāw=im bi-der-ewe*
 quarry=1SG SUBJ-give.PRS.2SG=ASP
 ‘Give me back my quarry’ (MacKenzie 1962: 8)
- b. *min žin=im nā-wē*
 1SG woman=1SG NEG-be.desirable.PRS.3SG
 ‘I don’t want a wife’ (MacKenzie 1962: 6)
- c. *heyās ū hesen memend=im be hīč le des*
 Heyas and Hasan Mamand=1SG for nothing from hand
der-čū-n
 PRV-go.PST-3PL
 ‘Heyas and Hasan Mamand have been lost **to me** for nothing’
 (MacKenzie 1962: 36)

It needs to be noted that the so-called clitic pronouns are not merely the reduced form of the corresponding free pronouns (see Table 4). In fact, only in functions A–C above can they be substituted by the corresponding form of the free pronouns. In function D, they are agreement markers, rather than pronouns, so the label is actually misleading in these contexts, though I will continue to use it as a cover term for members of the paradigm.

3.2.4.1. Clitic placement and displacement

The syntax of pronominal clitics is the most complex and theoretically challenging aspect of Sorani syntax. Considerations of space preclude a full treatment of this topic (see Öpengin 2016, in press, for detailed analysis of closely-related Mukri), but some of the more important principles will be briefly illustrated here. The complications involve almost exclusively clitics in the “Indirect Participant” function E above. In some cases, they are the complements of prepositions.

Because the pronouns concerned are clitics, their placement is determined by the interaction of distinct principles, including rules on clitic placement (e. g. sec-

ond-position), constraints on clitic stacking within a word, and faithfulness preferences that favour complements being adjacent to their heads.

In general, we expect the complements of adpositions to occur adjacent to the adposition concerned, as in *leget=tān* ‘with you(PL)’, i. e. respecting a basic faithfulness requirement. However, there are a number of contexts where pronominal clitics that are adpositional complements are required to move away from their syntactic heads and be hosted by another element. This is almost obligatory with some prepositions, when the clause concerned also includes an overt direct object, and it is not already hosting another argument-indexing clitic:⁴

- (19) čey su 'āl-ēk=it lē bi-ke-m
some question-INDF=2SG to.ABS SUBJ-do.PRS-1SG
'(that) I put some questions to you' (McCarus 2009: 617)

The preposition here is the multi-functional *le* ‘from, for, to’, which transforms into its so-called absolute form *lē* when its complement is a pronominal clitic. This rule affects certain prepositions in Sorani: *le* > *lē* ‘to, from, for’; *be* > *pē* ‘by, with’, =*(y)e* > *(y)ē* (remnant of a goal-flagging preposition, which now occurs solely as a clitic hosted by verbs when a goal argument immediately follows the verb; it is glossed DRCT in the examples). Thus compare *be min bi-tē* ‘say to me!’ (with an independent pronoun as complement of the preposition *be*) vs. *pē=m bi-tē!* (same meaning), with a pronominal clitic complement (MacKenzie 1961: 123). Returning now to (19), we find that the pronominal clitic is not realized locally on its preposition, but on the preceding direct object. Crucially, the preposition remains in its absolute form even when its pronominal complement is realized elsewhere.

Leftward displacement of prepositional complements is generally blocked in past-tense transitive constructions, because here the clitic pronoun which indexes the A takes precedence, and will be hosted by the direct object or other constituent of the VP (illustrated above), if available. If a past transitive construction also includes an adpositional argument, with a pronominal complement, then that pronominal complement is frequently displaced, but not to a preceding element. Instead, it is indexed via a verbal agreement affix. This is illustrated in (20), where the verbal suffix expresses the person value of the displaced complement of the preposition:

- (20) selām=yān lē kird-īn
greetings=3PL.A to.ABS do.PST-1PL
'They(A) greeted us', lit. 'did greetings for/to us' (McCarus 2009: 618)

⁴ The constraint on clitic stacking concerns clitics indexing (or bearing) a verbal argument; in the context of this rule, Indirect Participants act like verbal arguments. However, an argument indexing clitic can attach to a possessor-indexing clitic. There is also a general constraint against the repetition of phonetically identical pronominal clitics.

Not all adpositional phrases partake in this kind of construction; it is most widely-used with those adpositions governing what we have loosely termed an Indirect Participant, typically a benefactive, recipient, or external possessor.

If the adpositional complement is third person singular, the appropriate verbal agreement affix in the past tense is zero, so no overt indexing of the Indirect Participant occurs, except indirectly, through the absolute form of the preposition itself:

- (21) *kič [...] dergā=y lē kird=uwe*
 girl door=3SG.A for.ABS do.PST-PRF(3SG)
 ‘The girl(A) [...] opened the door for **him**’ (MacKenzie 1962: 30)

The prepositions *bō* ‘for’ and *legel* ‘with’ differ from those just discussed in that they allow the leftward displacement of a pronominal complement, but lack a corresponding absolute form:

- (22) *mes'ele=y ehme-y _bēyān=im bō b-ēn-ē*
 story-EZ Ahmed=EZ_carefree=1SG for SUBJ-bring.PRS-3SG
 ‘(That he may) bring the story of Ahmed the Carefree for **me**’ (MacKenzie 1962: 56)

3.2.5. Subordinate clauses

Sorani is consistently right-branching, with all types of subordinate clause following main clauses, and relative clauses following their heads. Like Kurdish generally, Sorani almost entirely lacks non-finite syntax. There are thus no infinitival constructions directly comparable to English complements in *-ing*, or *to+infinitive*. Modal verbs such as *wīstin* ‘be desirable, necessary’, and *twānīn* ‘be able’, take complement clauses with verbs in the present subjunctive, and no complementizer. Other complement-taking verbs, such as *zānīn* ‘know’ generally do not use a complementizer between main and complement clause. A complementizer *ke* is available for relative clauses, restrictive and unrestrictive, as in *selāhedīn*, *ke dīnye=y girt* ‘Saladin, who conquered the world’. Restrictive relative clauses may also be introduced through an Ezafe particle, on the condition that the head noun is marked as definite, either through the definiteness suffix *-eke*, or through a demonstrative determiner, or is itself pronominal (MacKenzie 1961: 132–133). In this case, they lack any complementizer; this is a further difference between Sorani and Northern Kurdish. Postposed relative clauses, on the other hand, require the complementizer *ke*, regardless of any other factors.

- (23) *ew-e=y tō dī-w-it-e*
 that-DEM=EZ 2SG see.PST-PRF=2SG.A-3SG
 ‘That which you have seen.’ (MacKenzie 1962: 133)

In Bingird and Piždar, non-restrictive relative clauses are introduced by *her ke*. In these dialects, the particle *agar* is used in a variety of clause types as a complementizer:

- (24) *de-zān-ē agar kuř=ī N. nī=ye*
 IND-know.PRS-3SG COMPL son=EZ (proper name) NEG=COP.3SG
 '(when he read it) he knows he is not the son of N.' (Bingird dialect,
 MacKenzie 1962: 132)

3.2.6. Alignment

As in many West Iranian languages, in Sorani the morphosyntax associated with past transitive verbs is distinct from that of other clause types in the language. This has often been referred to generally as “ergativity”, though it is not particularly meaningful to characterize an entire language in these terms, particularly Sorani; see Dabir Moghaddam (2012) for extensive discussion. In what follows, I will present the main outline of the system, without committing to any particular classification of the language as a whole.

There is no morphological case in Sorani, so the entire alignment discussion centres on patterns of agreement (or more neutrally, on indexing). The general pattern in Sorani is the obligatory indexing of S or A. In all environments except past transitive clauses, the person indexing is via one of the verbal suffixes shown in Table 6. In past transitives, however, the A is obligatorily indexed through the appropriate pronominal clitics provided in Table 8. In this function, the clitic is best considered an agreement marker, rather than any kind of pronoun (see Samvelian 2007; Haig 2008; Öpengin 2016). One argument in favour of this analysis is the presence of the clitic in constellations where a pronoun would not normally be expected, for example in sequences of same-subject clauses such as the following (glosses simplified), where a pronoun would normally be omitted in the second and third conjuncts:

- (25) a. *Kuř beyānī zū heštā*
 boy morning early rise.PST(3SG)
 b. *čūe lāy pāšā wut=ī*
 went side.of king say.PST(3SG)=3s.A
 ‘The boy rose early in the morning, went to the king, and said’
 (MacKenzie 1962: 56)

- (26) *šew pel=ī kuř=yān girt kird=yān=e*
 evening arm=EZ boy=3PL.A take.PST do.PST=3PL.A=DRCT
perde=we
curtain=ASP
 ‘In the evening they took the boy by the arm and put him behind the curtain’ (MacKenzie 1962: 520)

Perhaps even more telling is the presence of the A-clitic in subject relativization, where we would normally not expect any pronoun, as in the example mentioned above: *selāhedin, ke dinye=y girt* ‘Saladin, who conquered the world’. Unlike free pronouns, which are preferably dropped in such contexts (even in a pronoun-happy language like English), the clitic pronouns indexing a past-tense A cannot be omitted, but are required in all past tense transitive clauses. The other objection to considering them pronouns is that they occur in the presence of a coreferent subject NP. In short, the clitic pronouns in past-A functions are evidently better analysed as agreement markers, and most contemporary scholarship converges on this point.

The rules determining placement of the clitics have been discussed in §3.2.4.1 above. With regard to the verbal agreement affixes, with past tense transitive clauses they may index an object that is otherwise not present in the clause; otherwise the verb carries default third person agreement (zero). Alternatively, the agreement suffix may be co-opted to index an indirect participant; see (20) above.

3.2.6.1. Experiencers as Non-Canonical Subjects

The past transitive construction just discussed shows obvious similarities with clauses based on a set of (basically intransitive) experiencer-type predicates (e.g. of desire, physical perception and sensation, but also of possession). Here too the experiencer or possessor is obligatorily indexed via a pronominal clitic, adhering to similar clitic placement principles as those obtaining for past A clitics. Examples of this type of construction include the following:

- (27) *he=m=bū* ‘I had’ (existent=to.me=it.was)
- (28) *tīnū=m=e* ‘I am thirsty’ (thirst=to.me=it.is)

The verb *wīstin* ‘be desirable, necessary’ (present stem *-ewē-*) indexes the desirer through an obligatory pronominal clitic, and the ‘desired’ though a verbal affix, regardless of tense:

- (29) *e=y=ewē-m* ‘he wants me’ (IND=3SG=be.desired.PRS-1SG, MacKenzie 1961: 105)

In Wārmāwa dialect this construction is largely substituted by a similar construction based on non-verbal predicate, the Turkic loan *gerek*:

- (30) *gerek=im=e* ‘I want, need’ (*gerek*=1SG=COP.3SG)

The experiencers and possessors in this kind of construction are syntactically not distinct from other subjects, and exhibit the kinds of semantics widely associated with so-called Non-Canonical Subjects (cf. Onishi 2002) cross-linguistically, so it is reasonable to refer to the experiencers in constructions such as (28–30) as

Non-Canonical Subjects. The Non-Canonical Subjects associated with Central Kurdish experiencer and possessive predicates are most probably retentions of archaic structures in Iranian, and close parallels can be found in many contemporary West Iranian languages, though they have disappeared in the northernmost varieties of Kurmanjî Kurdish (Haig 2006; Haig and Öpentin 2018).

4. Northern Kurdish of North Iraq: Behdinî (Behd.)

The varieties of Northern Kurdish (NK) spoken in northern Iraq are referred to by their speakers as Behdinî. According to the classification of Öpentin and Haig (2014), Behdinî is part of the “Southeast” dialect group within Northern Kurdish, a group that spills over into neighbouring regions of Turkey such as Hakkari, and eastern Syria. Behdinî shares most of its morphosyntax with the better-known Northern Kurdish dialects of Turkey, and much of the description of NK found in Haig (this volume, chapter 2.3) also applies to Behdinî. In order to avoid repetition, this section focuses on those aspects of Behdinî that distinguish it from the rest of NK. My account draws largely on MacKenzie (1962) and Blau (1975), supplemented with information from a native speaker of Dohuk, currently living in Zakho.

4.1. Phonology

4.1.1. Consonants

Table 9: The consonant phonemes of Behdinî

	bilab.	lab.- dent.	dent.	alv.	post- alv.	pal.	vel.	uvul.	pharyn.	glot.
Plos.	p ^h p b			t ^h t t̪ d		k ^h k g	q			?
Fric.	v f	s z	ʂ ʐ	ʃ ʒ		x ɣ	ʁ	ħ ʕ ⁵	h	
Affr.				ɸ ɬ	dʒ					
Nas.	m		n			ŋ				
Trill			r							
Flap			f							
Approx.	w				j					
Lateral			l	(dialectally also l̪)						

⁵ As pointed out by Christiane Bulut (p.c.), in Kurdish as well as other languages of the region, these segments can be considered to be glottal stops produced with a retracted tongue root, rather than fricatives.

As discussed in Haig (this volume, chapter 2.3) the most notable feature of the Northern Kurdish consonants is the presence of an aspirated vs. non-aspirated distinction on the voiceless stops and affricates. However, MacKenzie (1961: 30) notes that the Akre dialect of Behd., and the varieties spoken by the surrounding Zēbārī tribal confederation, lack this additional series, bringing their systems closer to that of Sorani. MacKenzie does, however, include the pharyngeals /t̪ s̪ z̪/ in the consonant system of Behd., stating that they are identical in quality with the corresponding sounds of Arabic. They are retained in words of Arabic origin, such as *šebr* ‘patience’, or *teyr* ‘bird’, but also occur in the native vocabulary of Behd., e. g. *sār* ‘cold’, *tārī* ‘dark(ness)’, where they trigger backing of the following vowel (MacKenzie 1961: 36). Minimal pairs, or near-minimal pairs, demonstrating the phonemic status of the pharyngeals include *tā* ‘fever’ vs. *tā* ‘branch’, or *bez* ‘suet’ vs. *pez* ‘sheep’ (MacKenzie 1961: 35–36). Blau (1975: 28) does not include the pharyngeals in her analysis of the Sinjarī dialect. In general, the phonemic status of the pharyngeal consonants in Northern Kurdish remains a matter of debate, and although they are undoubtedly perceptually salient and constitute a feature of native pronunciation of the language, the functional load of pharyngeals remains limited in any variety of Northern Kurdish.

In more westerly varieties of NK, an inherited sequence [xw] is retained, or rendered as a labialized fricative[xʷ], while in Behd. it is generally reduced to a velar fricative [x] (Haig and Öpengin 2018).

Behd.	Standard Kurm.	Gloss
<i>xārin</i>	<i>xwārin</i>	‘to eat’
<i>xāndin</i>	<i>xwāndin</i>	‘to read, study’
<i>xo</i>	<i>xwe</i>	‘self’
<i>xē</i>	<i>xwē</i>	‘salt’

4.1.2. Vowels

The analysis of the Behd. vowel system is hampered by differences in the transcription conventions across different sources, and in standardized NK orthography. This presentation aims for a compromise solution between the widely-used orthography of NK, and the transcription of MacKenzie (1961, 1962). Table 10 gives an overview of the main differences, and establishes the symbols used in this description:

Table 10: Transcription of vowels for Behd.

IPA	MacKenzie (1961)	Stand. NK orthography	This chapter
[y]	û	(ü)	ü
[u:]	ô	û	ū
[o]	u	u	u
[o:]	(ō, for Šēxān dialect, 1961: 39)	o	o
[ɛ~æ~a]	a	e	e
[ɑ~ɔ]	ā	a	ā
[e:]	ē	ê	ē

All varieties of Behd. exhibit basically the same eight-vowel system that characterizes Northern Kurdish generally (see Haig, this volume, chapter 2.3). MacKenzie (1961: 33) assumes the system for the Akre dialect shown in Fig. 3.

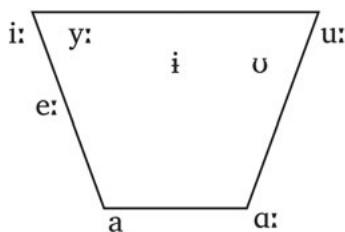


Figure 3: Vowel system of Behd., variety of Akre

Vowel length is not by itself distinctive, though [i, y, u, e, a] are generally pronounced longer and tenser than other vowels, particularly in open syllables. The main difference between Behd. and the rest of NK involve a set of changes, called here the Behd. Vowel Shift, and originally outlined in Blau (1975: 33). The changes affect the tense back rounded vowels:

1. The closed back rounded [u:] is fronted to [y:], and in many dialects, derounded to [i:], thus merging with inherited [i:]. For example, the standard Kurdish minimal pair [ʃu:r] ‘sword’ vs. [ʃi:r] ‘milk’ are both realized as [ʃi:r] in Dohuk and Zakho.
2. The close-mid back vowel [o:] raises towards [u]. Thus standard Kurmanjî [miro:v] ‘person’ is [miruv], and [go:t] is [gut] ‘said’ in Zakho. This latter process is also noted for Šemzinan Kurmanjî (Haig and Öpengin 2018), but it does not apply as regularly as the first one.

The full change [u: → i:] appears to be relatively recent. MacKenzie (1961), based on fieldwork from the 1950's, noted the fronting of [u:→ y] for Akre, and assumed [y:] to be the norm, but does not mention systematic derounding. He notes (1961: 40) that some speakers of Amediye and Zakho occasionally “confuse” [y:] with unrounded [i:], implying that derounding was at best sporadically attested at that time. Similarly, Blau (1975) based on fieldwork from the late 1960's, notes only fronting of [u:] for Amediye, but not derounding. More recently, Haig and Mustafa (2016) note the systematic shift of original [u:] to [i:] for most of the region surrounding Dohuk. It also occurs in loanwords such as [si:k] ‘market’ from Arabic *sug*. Zakho has apparently gone further than most other dialects in that derounding also applies to the third person singular past of ‘be’, which is [bu:] in Standard Kurdish, but in Zakho [bi:]. However, in no dialect is the past tense of ‘go’ affected, which generally contains [u:]. This is presumably because the [u:] here is a secondary development from [o:].

Haig and Mustafa (2016) also observe that the Ezidi speakers of the town of Šarya east of Dohuk retain [u:] (e. g. [bu:k], as opposed to [bi:k] ‘bride’ of the Sunnite speakers of Dohuk). This ties in with Blau's (1975) findings on the Ezidi speakers of Sinjar region, which likewise retain inherited [u:]. In the speech of an elderly male Ezidi speaker from Ba'adra, a predominantly Ezidi-inhabited township some 40km. from Dohuk, we also find no evidence of [i:] (e. g. the form for ‘everyone’ is [hæmu:], rather than Dohuk [hæmi:]). Thus while the Behd. vowel shift is still only poorly understood, it seems possible that it is utilized as a marker of religious boundaries; it certainly belongs to the kind of phonetic variants which Meyerhoff (2011: 26) labels as stereotypes, i. e. those that speakers are aware of, and can meta-linguistically comment on, hence are good candidates for markers of identity. However, the absence of the Behd. Vowel Shift among Ezidi speakers may also reflect distinct origins of the groups concerned; this remains to be investigated.

Table 11: Examples for the Behd. vowel shift

Gloss	Stand. Kurm.	Ezīdī of Šarya	Šemzinan	Dohuk	Zakho
‘hair’	<i>mu:</i>	<i>mu:</i>	<i>my:</i>	<i>mi:</i>	<i>mi:</i>
‘quick, soon’	<i>zu:</i>	<i>zu:</i>	<i>zy:</i>	<i>zi:</i>	<i>zi:</i>
‘bride’	<i>bu:k</i>	<i>bu:k</i>	<i>by:k</i>	<i>bi:k</i>	<i>bi:k</i>
‘sat’	<i>ru:nisti:</i>	<i>ru:nisti:</i>	<i>ry:nisti</i>	<i>ri:nisti:</i>	<i>ri:nisti:</i>
‘was’	<i>bu:</i>	<i>bu:</i>	<i>bu:</i>	<i>bu:</i>	<i>bi:</i>

Other differences from the dialects of Turkey include the lowering of [æ] towards [a], e. g. [bas] ‘enough’, and the backing of [a] to [ɑ] as in [bra:] ‘brother’.

4.2. Morphosyntax

Most of the features that set Behd. apart from other dialects of Northern Kurdish concern aspects of the predicate. Nominal morphology, on the other hand, is very close to Standard, or Botan, Kurmanjî. In the interest of brevity only those points are discussed where Behd. differs from the rest of Northern Kurdish (Haig, this volume, chapter 2.3). Some of the points mentioned are illustrated in the short text excerpt provided in §4.3 below.

4.2.1. Nominal morphology

The gender distinction in the singular is consistently observed in Behd., both in the form of the Oblique case: *-ī* (masc. sg.) vs. *-ē* (fem. sg.), and the forms of the Ezafe: *=yē* (masc. sg.) vs. *=yā* (fem. sg.). The oblique case of masculine singular nouns, often lacking or marked through Ablaut in other varieties of NK, is consistently marked on the noun via the suffix *-ī*: *min gut=e zelām-ī* ‘I(OBL) said=to the man-OBL’. The plural form of the Ezafe is *-ēt*, rather than the *-ē(n)* of most of NK. The paradigm of personal pronouns is identical to the rest of NK, except for the second person plural in Surčī and Akre, where the forms *engo* (Surčī, both direct and oblique), *hung* or *hing* (Akre, Amadiye, direct; (*hun*)*ngu*, (*hu*)*nge* oblique) are found.

4.2.2. Verbal morphology

In the person and number agreement paradigm, most varieties of Behd. resemble Sorani in that they distinguish the first person plural from the second and third person plural (-*īn* first person plural versus *-in* second and third person plural), though in Zakho this additional distinction is also reported absent, at least for some speakers. Elsewhere in NK, there is a single form *-in* for all persons in the plural. All varieties of Behd. preserve a historical final *-t* on the agreement suffix for the third person singular on the present tenses (MacKenzie 1961: 182). In Surčī, Amadiye, Akre and Dohuk the sufix is *-īt*, as in Dohuk *čē-nā-b-īt* PRV-NEG-be.PRS-3SG ‘It doesn’t happen’. In Zakho and Gullī we find *-it* (*t-č-it* IND-go. PRS-3SG ‘he goes’, Zakho, MacKenzie 1962: 362). Elsewhere in NK, the ending is reduced to *-e*. Thus we can recognize three grades of weakening of 3SG present: *-īt* (Dohuk), *-it* (Zakho), *-e* (most NK dialects of Turkey). Finally, most of Behd. has so-called ‘heavy’ present stems for the verbs *kirin* ‘do, make’ *dān* ‘give’, and *birin* ‘bring’ (*ke-*, *de-* and *be-* respectively), as opposed to *k-*, *d-* and *b-* in the rest of NK, though again, Zakho seems to pattern more like the rest of NK.

In certain contexts (generally subjunctive, or at least semantically irrealis), Behd. verbs take an ending *-itin* in the third person singular:

- (31) *zelām-ek ne hinde qewet jī dē šē-t wan dār-ā bi dest-ē xo hil-kēš-itin*
 man-INDF NEG that strength ADD FUT can.PRS-3SG
wan dār-ā bi dest-ē xo hil-kēš-itin
 those.OBL tree-PL.OBL with hand-EZ.M REFL PRV-pull.PRS-3SG
 '(even) a man not that strong will be able to pull those trees out with his hand' (Unger 2014)

The origin and function of this formative remain unclear. MacKenzie (1961: 91) notes an alternative “optional” third singular verbal suffix *-tin* in Arbil and Rewandiz dialects of Sorani, noting only that it occurs “most commonly in pause”. This is presumably related to the Behd. form, but does little to explain its function.⁶

The present indicative suffix, in NK generally of the form *di-*, is reduced in Behd. to [t-], and is transcribed as such here. The future tense in Behd. differs from the rest of NK in that the main verb lacks the subjunctive prefix *bi-*. Thus the future is created through a clause-initial *dē* (as in NK generally, cf. Haig, this volume, chapter 2.3, §2.3.4); this particle occurs generally in second position, but in the absence of an overt subject NP, may be clause initial) plus the inflected bare present stem of the verb. Functionally, the Behd. future has a much broader range of application than the label ‘future’ implies. It is frequently used for present habitual senses, as in the narrative provided in MacKenzie (1962: 361–362), where the speaker recounts the means of livelihood of his father (a charcoal-seller), alternating between simple indicative and future tenses to describe actions which evidently recur regularly, including in the past:

- (32) *I wērē dē dār-ā kom k-in*
 there FUT wood-PL.OBL gathering do.PRS-3PL
 ‘there (they) gather the wood’ (MacKenzie 1962: 362)

My impression is that the future tense is used with much greater frequency in Behd. than in the westerly varieties of NK, but this has not been systematically investigated.

Unique to Behd. is a past habitual, also used in a conditional sense, based on a particle *dā*, with similar positional constraints to future *dē*, plus the non-prefixes present stem. For example, an account of earlier traditional wedding celebrations (Haig and Mustafa 2016) begins with the following:

⁶ Geoffrey Khan (p.c.) points out that in NENA dialects in the Behdini area a *-ni* or *-in* augment is often added to verbal endings (of all persons) in pause, e. g. *garəš* ‘he pulls’ > *garəšni* (pause), *garšet* ‘you (ms.) pull’ > *garšetin* (pause). There is no clear Semitic etymology for this. There may be a link here, but the Behd. *-in* is restricted to third persons only.

- (33) *zelām dā hē-n* [...]
 man COND come.PRS-3PL
 ‘the men would / used to come [...]’

This form is also used for irrealis forms with non-past semantics; see Unger (2014). A further feature distinguishing Behd. from the rest of NK is the very frequent use of the aspectual clitic *-ewe* ‘back, again’, presumably reflecting the areal proximity to Sorani.

The final feature discussed here is likewise unique to Behd. within NK, and indeed within the entirety of Kurdish. Behd. has developed a present continuous, based on the Ezafe particle (indicating gender and number of the subject) and the present indicative form of the verb. Example (34) contrasts the use of this form, indicating immediacy of the event, with the simple indicative present (35), where a general present tense is implied:

- (34) *tu yē cotī t-ke-yī*
 2SG EZ.M ploughing IND-do.PRS-2SG
 ‘Are you ploughing? (right now)’

- (35) *ev čend sal=e tu cotyarī-yē lē*
 this how_many year=COP.3SG 2SG cultivation-OBL.F on.it
t-ke-yī
 IND-do.PRS-2SG
 ‘How many years is it that you are cultivating it?’

This usage is presumably related to the use of the Ezafe with non-verbal predicates, particularly with expressions of location, and with participles, described in MacKenzie (1961: 205–208), and Haig (2011). Finally, as in Sorani, Behd. makes extensive use of non-canonical subjects (in the oblique case) with predicates of possession, desire, and some physical sensations (heat, cold, but not hunger and thirst). The following example illustrates an oblique subject of a predicate of perception;⁷ note also in this example the use of the ezafe particle following the oblique subject, which is typical for non-verbal predicates expressing temporary states in Behd.:

⁷ In Dohuk, to express ‘hear’ a different construction is used, involving a complex predicate consisting of *gūlē* combined with the light verb *būn*. In the present tense, the subject is treated canonically (i. e. is in the direct case). Thus ‘I (female) hear the sound of the children’ is: *ez a dengē pičīkā gūlē dibim* (Baydaa Mustafa, p.c.). Remarkably, in the past tense the subject goes into the oblique case, indicating that this expression is treated as overall transitive, despite the intransitive light verb *būn* ‘be’.

- (36) *min=ē gū lē*
 1SG.OBL=EZ.M.S ear at.it
 ‘I hear it’ (spoken by a male speaker, hence the masc. sg. form of the Ezafe, MacKenzie (1961: 206), dialect of Akre)

4.3. Short text in Behdinî Kurdish

The following text is an excerpt from a traditional tale, told by a male Ezidi speaker, approximately 60 years old at the time. The text was recorded in 2011 during fieldwork in the township of Ba’adra, undertaken by the author together with Ergin Öpengin.

- (37) *der-ā mirin lē he-b-itin*
 place-EZ.F death in.it existent-COP.PRS.SUBJ-3SG
ez lē nā-b-im
 1SG in.it NEG-COP.PRS-1SG
 ‘(He says:) I won’t stay at a place where death (mortality) is’
- (38) *hēj ko ne-čū māl-ā xo*
 before NEG-go.PST(3SG) house-EZ.F REFL
 ‘before he went back to his home’
- (39) *ū išāret-ek bo xo dā-nā [...]*
 and sign-INDF for REFL PRV-put.PST(3SG)
 ‘and placed a sign for himself’
- (40) *dā b-zān-im išāret-ek wā he=ye*
 COND SUBJ-know.PRS-1SG sign-INDF such existent=COP.3SG
 ‘(saying) that I know there is such a sign’
- (41) *wāxtē čū hevāl ū bend-ēt xo helā-n⁸*
 when go.PST(3SG) for and comrade-EZ.PL REFL leave.PST-3PL
 ‘when he departed and left his friends and comrades’ [...]
- (42) *cāmēr-ek-ī got=ē te xēr=e*
 man-INDF-OBL.M say.PST(3SG)=to.him 2SG.OBL goodness=COP.3SG
bāb-o
 father-VOC.M
 ‘a man said to him: “what are you doing, old fellow?”’

⁸ The verb is *helān* ‘leave’, which in other dialects has the infinitive *hištin*. Note the regular number agreement with the object, typical of the canonical ergative construction in Behd.

- (43) *dē kēwe č-ī?*
 FUT where go.PRS-2SG
 ‘where are you going?’ (note the use of the future tense)
- (44) *go bi xādē ez dē č-im=e*
 say.PST(3SG) by God I FUT go.PRS-1SG=DRCT
Qub-ā Felek-ē
 dome-EZ.F Heaven-OBL.F
 ‘he said: “By God, I am going to Heaven’s Dome” [...]’
- (45) *yā’nī dē č-īt=e Qub-ā Felek-ē*
 so FUT go.PRS-3SG-DRCT dome-EZ.F Heaven-OBL.F
 ‘so that is, he would go to Heaven’s Dome’

5. Varieties of Gorani in northern Iraq

5.1. Overview

Following Mahmoudveysi et al. (2012), I use Gorani (spelled *Gūranī* in Bailey 2016, and *Gurāni* in MacKenzie 2002) as a cover term for a group of West Iranian languages, with a probable historical epicentre in the mountainous Hawromān region of western Iran. The best-described, and morphologically most complex, variety of Gorani is Hawrāmī, which has a reasonably well-established written standard, based on the variety of Pāve in Iran. The historical forerunners of Gorani spilled westward from their mountainous origins into today’s Iraq, and are currently still spoken by a geographically scattered and socio-culturally diverse group of speech communities, beginning in the Iraq-Iran border region between Halabja and Xanaqīn, and extending northwest towards Mosul. According to MacKenzie (2002), the original speech zone of Gorani in this region was much larger, and today’s scattered Gorani islands thus represent the remnants of what must have been a more extensive contiguous region.

The most reliable and accessible description of any variety of Hawrāmī remains MacKenzie (1966). Recently, documentation and analysis of local varieties of Gorani spoken in Iran have been published as part of language documentation projects (Mahmoudveysi et al. 2012; Mahmoudveysi and Bailey 2013; Bailey 2016). The varieties of Gorani spoken in Iraq, on the other hand, remain very poorly documented. Bailey (2016) summarizes much of the existing scholarship on Gorani, and this section largely adopts her conventions. Bailey identifies the following varieties as belonging to Gorani (see the map in Fig. 4 for locations).

Hawrāmī

As mentioned, this is the name given to those varieties spoken in the mountainous Hawrāmān (Awromān) regions of western Iran and eastern Iraq. It is commonly differentiated into Hawrāmān-i Taxt, the varieties of the high mountain regions, and Hawrāmān-i Luhon, the region of valleys. Varieties of Hawrāmī are spoken in Iraq in and around the city of Halabja, close to the Iranian border. They are not treated in this chapter; see Mahmoudveysi and Bailey, this volume, chapter 4.5.

Kandūlayī

Kandūlayī refers to the variety of Gorani spoken in a complex of three villages some 50 km. north of Kermānshāh (see Fig. 4). This variety is close to Hawrāmī of Pave.

Zardayāna

The variety spoken in the village of Zarda, spoken in a village about 100 km. north of Sar Pol-e Zahāb in Iran. It is documented in Mahmoudveysi and Bailey (2013).

Gawrājūyī

The variety spoken only in the village of Gawrāju, located about 10 km. west of the township of Gahvāre in West Iran. It has been described in Mahmoudveysi et al. (2012), and Bailey (2016). The dialect of Gawrāju is relatively isolated from the rest of Gorani, and has been heavily influenced by surrounding varieties of Southern Kurdish. This may explain why it has lost many of the identifying features of Gorani (see below).

Bājałānī

This is a term loosely identified with the speech of the Bājałān tribes, spoken in several locations just east of Mosul, but also villages near Xanaqīn and Koy Sanjak to the southeast near the Iranian border. The Bājałān are sometimes subsumed under the Šabak (see below), but I follow Bailey (2016), basically adhering to MacKenzie (1956), in maintaining a difference between the two groups, which seems to be more in line with their self-perceptions. Along with the language (here considered a variety of Gorani), the Bājałān are known for their religious heterodoxy.

Šabakī

The speakers are socially and linguistically closely aligned with the Bājałān, and are originally located east of Mosul on the Nineveh plains. The claim that Šabakī is a mixture of Turkish, Arabic and Kurdish is not substantiated by the available data, which suggest that it is a variety of Gorani, very close to Bājałānī (Bailey 2016: 643).

Kākayī

The name Kākayī is generally applied to groups belonging to the Yaresan or Ahl-e Haqq religious community, around the towns of Tōpzāwa near Kirkūk, Xānaqīn and Arbil (Bailey 2016: 644). The language is often referred to with the term *Mačo* (3sg present of the verb ‘say’).

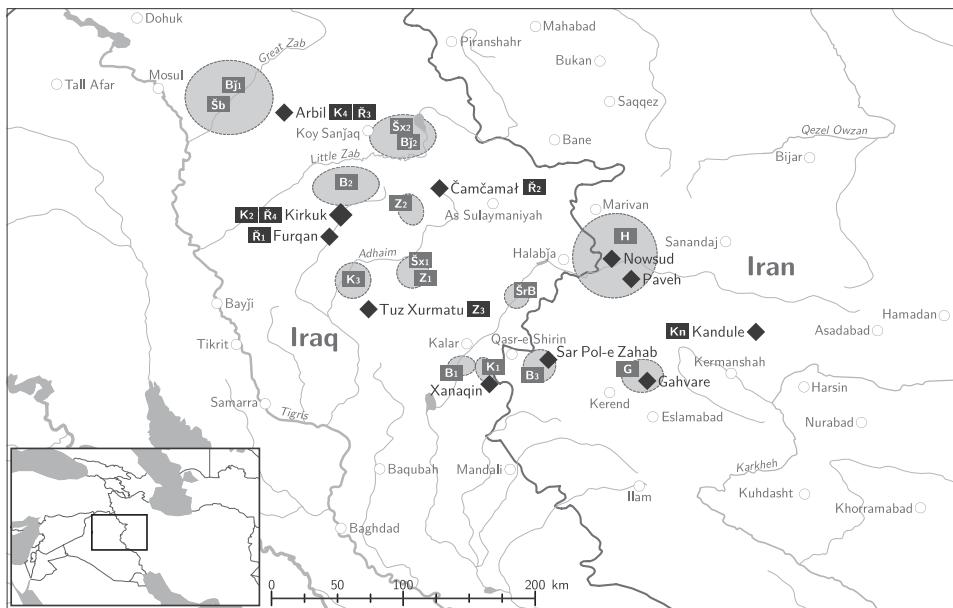


Figure 4: Names and locations of Gorani speaking localities
(adapted from Bailey 2016: 9–10)

Key to Abbreviations used in Fig. 4:

B (Bēwyānī) B1: Sarqizil, Bardī ‘Alī Xwārū/Žūrū, Say Mişafā, Gōrī Ginūž, Kānī Šīrna; B2: Bēwyānī Gawra/Bičūk and about another fifteen villages; B3: Bēwyān, Dūšamnān (their main places in this area), Sar Pol-e Zahāb, Dāraka, Qalama, Šāy Tōtyā, Barxu Bārānī Xwārū/Žūrū, Mijūryānī Ambar/Awbar, Tangī Ḥamām, Sarqałā

Bj (Bājalānī) Bj1: several villages near Al-Mawṣil (members of the tribe are also present in Xānaqīn, Kalār, etc.); Bj2: Kānī Māz, Taqtaq, Sē Girdkān, Qāmīš, (Kōya (Koy Sanjaq), Dukān)

G: Zarda and Gawrajū villages

H (Hawrāmān area) Center: Pāve and Nowsūd

K (Kākayī) K1: Xānaqīn, Mēxās, Qarāmān, Malā Ṙahmān, Dārā, Qalama, ‘Alī Bāpīr, Ṙamazān, Tapa Čarmē, Ḥājī Mişafā; K2: Kirkūk (districts Ḥayy ‘Askarī, Usarā’ al-Mafqūdīn etc.); K3: around Tōpzāwa (center of the Kākayī), ‘Alī Sarāy, Ġarja Kōyī, Ma’tīq, Dāquq, Zaqr; K4: near Arbīl: Sufaya, Wardak, Kabarlū, Tūlaban, Gazakān

Kn (Kandūla) three villages

Ŕ (Ŕožbayānī) Ŕ1: Furqān; Ŕ2: Čamčamāl; Ŕ3: Arbīl; Ŕ4: Kirkūk and Laylān (near Kirkūk)

Şb (Şabak) Şabak villages near Al-Mawṣil

ŞrB (Şaraf Bayānī) villages around Bamō

Şx (Part of Şexānī) Şx1: villages near Qādir Karam: Qaşqa, Wēla, Şawak, etc.; Şx2: Kānī Māz, Taqtaq, Sē Girdkān, Qāmīš, (Koya (Koy Sanjaq), Dukān)

Z (Zangana) Z1: Qādir Karam; Z2: Bakragaṛa, Sipasar and many more (in Xānaqīn, Kalār, Kifrī (between Kalār and Tūz Xurmātū), Arbīl and surrounding area and in the surroundings of Al-Mawṣil); Z3: Tūz Xurmātū

5.2. Gorani and Kurdish

The relationship of Gorani, as delineated above, to Kurdish is quite complex. Most speakers refer more or less collectively to all the Iranian languages of Iraq as some form of *Kurdī*, and Leezenberg (1994: 3) states that the Gorani speakers he met in Iraq all considered themselves “Kurds”. Unsurprisingly, centuries of close contact between speakers of Gorani and speakers of Central (and in some places Northern) Kurdish have yielded convergence in phonology, but also in syntax and lexicon. MacKenzie (2002) considers the phonology of Gorani basically identical to that of Sorani (cf. §3.1). Nevertheless, from a historical Iranianist perspective there are good reasons for distinguishing Gorani from Central and Southern Kurdish, and most scholars of Iranian languages continue to do so.

Verb paradigms are among the more reliable means of assessing historical relatedness. Sample paradigms of the present indicative are provided in Table 12. For Hawrāmī of Luhon (Iran, MacKenzie 1966: 37) and Bājālānī (Şabakī, according to the speaker) of the village of Arpačī, a few kilometres east of Mosul (MacKenzie 1956: 424) the verb is ‘sleep’ and for “Mačo” of Topzawa (Leezenberg 1994), the verb is ‘see’. The corresponding person agreement suffixes from Zardayāna of West Iran (Mahmoudveysi and Bailey 2013: 43) are also provided. As a comparison with Kurdish, the corresponding present-tense suffixes from Sorani Kurdish (Table 5 above) are included.

Table 12: Present indicative verb paradigms in Gorani

	GORANI			KURDISH	
	HAWR. 'sleep'	BAJ./ ŠABAKĪ 'sleep'	MAČO 'see'	ZARD.	SORANI
1SG	<i>m-us-u</i>	<i>m-ōs-ī</i>	<i>me-wīn-ū</i> [mōjny]	-ī/y	-im
2SG	<i>m-us-ī</i>	<i>m-ōs-ī</i>	<i>me-wīn-ū</i>	-ī	-īt, -yt
3SG	<i>m-us-o</i>	<i>m-ōs-ō</i>	<i>me-wīn-o</i>	-o	-ē(t)
1PL	<i>m-us-me</i>	<i>m-ōs-mē</i>	<i>me-wīn-im</i>	-mē	-īn
2PL	<i>m-us-de</i>	<i>m-ōs-ē</i>	<i>me-wīn-de</i>	-dē	-in
3PL	<i>m-us-ā</i>	<i>m-ōs-ān</i>	<i>me-wīn-ā</i>	-ān, -ānē	-in

A number of points of interest emerge from this comparison. First, the syncretism of first and second singular forms in all varieties of Gorani outside of Hawr. Second, the phonological shape of the person markers in Gorani, which are both relatively internally consistent, and exhibit obvious differences to the corresponding paradigms from Sorani.

Table 13 below provides a summary of ten morphological features common to most of Gorani, but lacking in Kurdish. The features are selected from the extensive comparative tables provided in Bailey (2016: 648–668), and include only those features that (i) are present in at least three of the varieties identified as Gorani; and (ii) are lacking in Northern, Central and Southern Kurdish. Thus features such as the clitic =*ewe* ‘back, again’, which is characteristic of all Iranian languages of Iraq (and has even spread to non-Iranian languages), are not included. Likewise, a feature such as a first person plural pronominal clitic with the form =*mā(n)*, with cognates throughout much of West Iranian, are of little value in distinguishing Gorani from Kurdish, and are hence not included.

Table 13 is far from exhaustive, and should only be considered as a tentative and partial set of candidates for morphological innovations that set Gorani apart from its Kurdish relatives. Not all features are of equal significance. The presence of a feature such as the -*g(i)n-* present stem of ‘fall’ in both Hawrāmi, and geographically quite distant Bājalānī, is highly significant; such a feature is hardly likely to have been independently borrowed, nor does it seem likely that all intervening languages had this feature, and then dropped it. It would therefore be more likely to consider it an inheritance from a presumed common ancestor of these two languages, i. e. some (branch of) proto-Gorani. It will be noted that Gawrājūyī (abbreviated Gawr.) differs from the rest of Gorani on several counts, mostly exhibiting the corresponding forms from Southern Kurdish. This confirms the overall impression that Gawr. is the most heavily Kurdisized variety of Gorani yet documented. A question mark indicates lack of relevant data; I have

collapsed Bailey's two columns “Bājālānī” and “Bājālānī/Šabakī” into a single column.

Table 13: Lexical and morphological features of Gorani not attested in Kurdish

	Gawr.	Haw.	Kand.	Zard.	Kāk.	Baj./Šab.
At least one demonstrative pronoun stem contains <i>-ēd-</i> (Bailey 2016: 651, 654)	–	+	–	+	?	+
Reflexive pronoun stem <i>w-</i> (Bailey 2016: 655) ⁹	– (<i>ištān</i>)	+	+	+	+	<i>hē</i> (<* <i>wē?</i> , cf. <i>hūn</i> 'blood')
3SG present indic. verbal suffix with a back rounded vowel (Bailey 2016: 657)	rare	+	+	+	+	+
2PL imperative contains a <i>-d-</i> (Bailey 2016: 659)	–	+	+	+	?	–
unstressed present indicative prefix <i>m+</i> unrounded, central/open, vowel (e. g. <i>m(i)-</i> (Haw., MacKenzie 1966: 32), or <i>ma-</i> elsewhere (Bailey 2016: 656). ¹⁰	+	+	+	+	+	+

⁹ According to Sara Belelli (p.c. 17. 09. 2016), a reflexive pronoun *wiž* (possible variant *viž*) is present in most ‘Laki-Kermānshāhi’ varieties (e. g. Bisotun, Chehr, Harsin, Pāyrvand: see Fattah, 2000: 291 and Belelli, 2016: 65–66). Depending on how one defines “Southern Kurdish”, this feature may not qualify. It is also attested in varieties of Lekī.

¹⁰ One might conjecture on possible influence of the Persian *mī-* prefix here. Two facts militate against this assumption: (i) the considerable differences in vowel quality and stress patterns (evident in the assimilation of the prefix vowel to stem-initial vowels in Gorani); (ii) the histories of the speech communities: in order for bound verbal morphology to be affected in this way, there would have to have been prolonged heavy contact and bilingualism with Persian across the entirety of Gorani. This seems unlikely given the locations and history of the Gorani speech communities. Sara Belelli (p.c. 17. 09. 2016) notes the presence of *m(a)-* (normally accompanied by a clitic =*a* attaching to the element immediately preceding the verbal form) in most ‘Laki-Kermānshāhi’ dialects (e. g. Bisotun, Chehr, Harsin, Pāyrvand). This morpheme is also typical of Laki dialects (see, for instance, Lazar, 1992: 218); see Fattah (2000: 371–372), Belelli (2016: 99–100). Thus the reservations that apply to the reflexive feature also apply here.

	Gawr.	Haw.	Kand.	Zard.	Kāk.	Baj./Šab.
1SG past intransitive verbal suffixes contain <i>-n-</i> (Bailey 2016: 658)	- (<i>-im</i>)	+	+	+	+	- (<i>-im</i>)
Third person clitic copula, present tenses contain <i>-n-</i> (Bailey 2016: 661) ¹¹	+	+	+	+	+	+
‘what’+ <i>-š</i> (the interrogative word ‘what’ carries a final sibilant, e. g. Zard. <i>čiš</i> , Bailey 2016: 665)	-	+	+	+	+	+
- <i>g(i)n-</i> stem for present tense of ‘fall’ (e. g. <i>magnō</i> ‘it will fall’, Baj., MacKenzie 1956: 422)	-	+	+ (<i>mangu</i>)	-	-	+
3SG present indic. ‘say’: <i>māč+o/ū</i>	-	+	+	+	+	+

Obviously we are still far short of a true comparative study of Gorani and Kurdish, but these few facts should suffice to demonstrate that despite the scattered and isolated nature of the Gorani speech communities in northern Iraq, they have retained a core of common morphological and lexical features that distinguish them from the Kurdish spoken by the surrounding speech communities.

5.3. Notes on syntax

The variety of Hawrāmī spoken in and around Halabja is very close to that of Pāve, thus has retained gender and case on nouns. Elsewhere, however, gender appears to have been lost. For the variety of Baj. described in MacKenzie (1956), an oblique case is apparently maintained:

- (46) *ī zelām-a-y činī hē=t b-er-e*
 this man-DEM-OBL with REFL=2SG.POSS SUBJ-bring.PRS.IMP-2SG
 ‘Bring this man with you!’ (MacKenzie 1956: 426, glosses added)

However, no other clear example of an oblique-marked direct object was available in MacKenzie’s material, which makes this example doubtful. Similarly, no evidence for an oblique case can be found in the scanty material of Leezenberg (1994). The following example is provided by Leezenberg from Mačo of Topzawa

¹¹ Along with a copula in =*ān*, Gawr. also has a 3sg copula -*ē*, which is actually attested much more frequently than =*ān* (261 tokens of =*ēn* versus 15 tokens of =*ān* in Bailey’s Gawr. corpus, Denise Bailey p.c.). This could be interpreted as further indication of the heavy Kurdification of Gawr.

(unfortunately all the examples of direct objects in Leezenberg's material are indefinite, so the possible effects of DOM cannot be ruled out):

- (47) *Ew zilām-ē me-wīn-ō*
 3SG man-INDF IND-see.PRS-3SG
 'He sees a man'

All varieties of Gorani use pronominal clitics to index an A in past transitive constructions, with the A-clitic hosted by the first constituent of the VP, as described for Sorani in §3.2.4:

- (48) *etū čēš=it wāt bene=šān*
 2SG what=2SG.A say.PST(3S) to=2PL
 'What did you say to them?' (Baj., MacKenzie 1956: 431)

Unlike Central Kurdish, the third person singular pronominal clitic is *=(i)š* in Gorani; otherwise the paradigm of pronominal clitics is cognate with the forms given in Table 8 above.

Abbreviations

ABS	Absolutive	INDF	Indefinite
ADD	Additive	M	Masculine
ASP	Aspect	MOD	Modal
CEZ	Compound Ezafe	NEG	Negation
COMPL	Complementizer	OBL	Oblique
COND	Conditional	PL	Plural
COP	Copula	POSS	Possessive
DEF	Definite	PRF	Perfect
DEM	Demonstrative	PRS	Present
DRCT	Directional particle	PRV	Preverbal particle
EZ	Ezafe	PST	Past
F	Female	REFL	Reflexive
FUT	Future	SG	Singular
IMP	Imperative	SUBJ	Subjunctive
IND	Indicative	VOC	Vocative

References

- Bailey, Denise. 2016. *A grammar of Gawrajū Gūrānī*. Göttingen: Universität Göttingen dissertation.
- Belelli, Sara. 2016. *A study on language and folklore in the city of Harsin (in Kermānshāh province of West Iran. Sketch grammar with texts and lexicon*. Naples: L'Orientale University PhD dissertation.

- Blau, Joyce. 1975. *Le Kurde de 'Amādiya et de Djabal Sindjār: Analyse Linguistique, Textes Folkloriques, Glossaires*. Paris: Klincksieck.
- Fattah, Ismaïl Kamandā. 2000. *Les dialectes kurdes méridionaux. Étude linguistique et dialectologique*. Leuven: Peeters.
- Haig, Geoffrey. 2006. Turkish influence on Kurmanji: Evidence from the Tunceli dialect. In Lars Johanson & Christiane Bulut (eds.), *Turkic-Iranian contact areas: Historical and linguistic aspects*, 283–299. Wiesbaden: Harrassowitz.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A Construction Grammar approach*. Berlin & New York: Mouton De Gruyter.
- Haig, Geoffrey. 2011. Linker, Relativizer, Nominalizer, Tense-Particle: On the Ezafe in West Iranian. In Foong Yap, Karen Grunow-Hårska & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives, vol. 1, Sino-Tibetan and Iranian Languages*, 363–392. Amsterdam & Philadelphia: Benjamins.
- Haig, Geoffrey. 2017. Western Asia: East Anatolia as a transition zone. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 396–423. Cambridge: Cambridge University Press.
- Haig, Geoffrey. In press. Debonding of inflectional morphology in Kurdish and beyond. In Songül Gündoğdu, Ergin Öpentin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*. Bamberg: Bamberg University Press.
- Haig, Geoffrey & Baydaa Mustafa. In press. Language choice and patterns of usage among Kurdish speakers of Dohuk: An empirical intergenerational study. In Songül Gündoğdu, Ergin Öpentin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*. Bamberg: Bamberg University Press.
- Haig, Geoffrey & Ergin Öpentin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 3. 99–122.
- Haig, Geoffrey & Ergin Öpentin. 2018. Kurdish in Turkey: Structure, variety, and status. In Christiane Bulut (ed.), *Linguistic minorities in Turkey and Turkic speaking minorities of the peripheries*, 157–230. Wiesbaden: Harrassowitz.
- Hassanpour, Amir. 1992. *Nationalism and language in Kurdistan 1918–1985*. San Francisco: Mellen Research University Press.
- Jahani, Carina. 2015. On the definite marker in Modern Spoken Persian. Paper held at the *Sixth International Conference on Iranian Linguistics* (ICIL6), Ilia State University, G. Tsereteli Institute of Oriental Studies, Tbilisi / Georgia, 23–26 June, 2015.
- Lazard, Gilbert. 1992. Le dialecte laki d'Aleshtar (kurde méridional). *Studia Iranica* 21. 215–245.
- Leezenberg, Michael. 1994. The Shabak and the Kakais: Dynamics of ethnicity in Iraqi Kurdistan. *ILLC Technical Reports Series X*. Amsterdam: Institute for Language, Logic, Language, and Computation.
- MacKenzie, David. 1956. Bājālānī. *Bulletin of the School of Oriental and African Studies* 18(3). 418–435.
- MacKenzie, David. 1961. *Kurdish dialect studies I*. Oxford: Oxford University Press.
- MacKenzie, David. 1962. *Kurdish dialect studies II*. Oxford: Oxford University Press.
- MacKenzie, David. 1966. *The dialect of Awroman (Hawrāmān-ī Luhōn). Grammatical sketch, texts, and vocabulary*. København: Det Kongelige Danske Videnskabernes Selskab.
- MacKenzie, David. 2002. Gurāni. *Encyclopædia Iranica* 6.4, 401–403. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/gurani> (accessed 28 August 2016).

- Mahmoudveysi, Parvin, Denise Bailey, Ludwig Paul & Geoffrey Haig. 2012. *The Gorani Language of Gawrajū, a village of West Iran. Texts, grammar and lexicon*. Wiesbaden: Reichert.
- Mahmoudveysi, Parvin & Denise Bailey. 2013. *The Gorani Language of Zarda, a village of West Iran. Texts, grammar and lexicon*. Wiesbaden: Reichert.
- McCarus, Ernest N. 2009. Kurdish. In Gernot Windfuhr (ed.), *The Iranian Languages*, 587–633. London & New York: Routledge.
- McDowall, David. 2004. *A modern history of the Kurds*, 3rd edn. London: Tauris.
- Öpentin, Ergin. 2016. *The Mukri variety of Central Kurdish. Grammar and Texts*. Wiesbaden: Reichert.
- Öpentin, Ergin. In press. Accounting for the combination of clitic and affix person markers in Central Kurdish. In Songül Gündoğdu, Ergin Öpentin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*. Bamberg: Bamberg University Press.
- Öpentin, Ergin & Geoffrey Haig. 2014. Regional variation in Kurmanjî: A preliminary classification of dialects. *Kurdish Studies* 3. 143–176.
- Samvelian, Pollet. 2007. A lexical account of Sorani Kurdish prepositions. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 235–249. Stanford: CSLI Publications.
- Skrzypek, Dominika. 2010. Between a demonstrative and an article. The status of -in in Old Swedish. *Folia Scandinavica* 11. 145–162.
- Unger, Christoph. 2014. Badini Kurdish modal particles dê and da: Procedural semantics and language variation. *Kurdish studies* 3. 177–200.

3.4. The Neo-Aramaic dialects of northern Iraq

Geoffrey Khan

1. Genetic affiliation

The Neo-Aramaic dialects of northern Iraq are modern vernacular forms of Aramaic, a Semitic language, which has a documented history in the Middle East of over 3,000 years, the earliest inscriptions being datable to approximately 1,000 B.C.E. The dialects in Iraq belong to the North-Eastern Neo-Aramaic (NENA) subgroup of Neo-Aramaic.¹ This is a highly diverse subgroup of over 150 dialects spoken by Jews and Christians originating from towns and villages east of the Tigris River in northern Iraq, southeastern Turkey and western Iran. The NENA subgroup is distinct from three other subgroups of Neo-Aramaic. These include the western subgroup spoken by Christians and Muslims in the villages of Ma'lula, Bax'a and Jub'adin in the region of Damascus, the Turoyo subgroup, spoken by Christians in the Tür 'Abdīn region of southeastern Turkey (see Khan, this volume, chapter 2.5), and Mandaic, spoken by Mandaeans in the cities of Ahwaz and Khorramshahr in Iran (see Khan, this volume, chapter 4.4). None of these subgroups is as diverse as NENA.

The NENA dialects are not direct descendants of any of the earlier literary forms of Aramaic, although they exhibit close affinities to Syriac and Jewish Babylonian Aramaic. The dialects rather have their roots in a vernacular form of Aramaic that existed in antiquity in the region of northern Mesopotamia, which differed from the vernacular underlying the literary languages of Syriac to the west and Jewish Babylonian Aramaic to the south. This is shown by the fact that, although exhibiting numerous innovations, they are more conservative than Syriac and Jewish Babylonian Aramaic in some features (Khan 2007; Fox 2008). Some of the dialects, moreover, have preserved lexical items of apparently Akkadian origin that do not appear in dictionaries of the earlier forms of literary Aramaic.² Structural differences among the NENA dialects are likely to reflect, to some extent, migrations of communities in the northern Mesopotamian region.

Within NENA itself one may identify a number of subgroups on the basis of linguistic structure and lexicon. There is a fundamental split between the dialects spoken by the Jews and those spoken by the Christians. This applies even to cases where Jewish and Christian communities lived in the same town, such as Koy Sanjak and Sulemaniyya.

¹ The term was coined by Hobermann (1988: 557).

² For Christian dialects see Krotkoff (1985) and Khan (2002a: 515, 2017: 262), and for Jewish dialects Sabar (2002: 12).

Within Jewish NENA dialects of Iraq two main subgroups are clearly identifiable: (1) the so-called *lišana deni* subgroup, which was spoken in the northwest of Iraq mainly in Dohuk province in locations to the west of the Great Zab river, such as Zakho, Dohuk, Amedia, Betanure, and Nerwa. (2) Dialects spoken in locations east of the Great Zab river in the Arbīl and Sulemaniyya provinces, e. g. Rustaqa, Ruwanduz, Koy Sanjak, villages of the plain of Arbel (Irbīl),³ the village of Dobe which is on the western bank of the Great Zab, Halabja and Sulemaniyya to the east, and as far south as Khanaqin on the Iranian border. This subgroup is generally referred to as trans-Zab (following Mutzafi 2008b). In addition, there was a small cluster of dialects in the region of Barzan, located in Iraq between these two areas, which exhibit a linguistic profile that is transitional between the two main subgroups (Mutzafi 2002, 2004b). The Jewish trans-Zab subgroup continues across the border into Iran (see Khan, this volume, chapters 2.5 and 4.4).

The divisions among the Christian NENA dialects of Iraq on structural and lexical grounds are not so clear-cut. One may, nevertheless, identify subgroups of dialects with distinctive features. These include the following: (1) the dialects of the Mosul plain in Nineveh province (e. g. Qaraqosh, Alqosh, Telkepe, Barətla), (2) the dialects in the far north close to the Turkish border in Dohuk province mainly west of the Zab (e. g. Aradhin, Barwar [= Barwari Bala], Nerwa, Derigne) and (3) the dialects east of the Zab in Arbīl and Sulemaniyya provinces (e. g. Ankawa, Shaqlawa, Bədyl, Koy Sanjak, Sulemaniyya). As can be seen, subgroups (2) and (3) correspond broadly to the Jewish *lišana deni* and trans-Zab subgroups respectively. In southeastern Turkey several dialects were spoken that are closely related to the dialects of the Iraqi Christian subgroup (2), especially in the Tiyare region (Khan, this volume, chapter 2.5). Subgroup (3) has one outlying dialect in western Iran (Sanandaj) (Khan, this volume, chapter 4.4). These three Christian subgroups in Iraq are distinguished for convenience of presentation,⁴ but it must be stressed again that this division is very broad-brush. The subgroups exhibit considerable internal diversity and the boundaries between them are not always clear. Fig. 1 and Fig. 2 below show the locations of the Jewish NENA dialects and the Christian NENA dialects respectively. In the following presentation the Iraqi Jewish subgroups are referred to as IJ1 (= *lišana deni*) and IJ2 (= trans-Zab), and the Iraqi Christian subgroups as IC1, IC2 and IC3.

³ The Jews in the town of Arbel itself spoke Arabic (Jastrow 1990).

⁴ For other classifications see Socin (1882: v), Duval (1896: 125), Maclean (1901: ix–xi) and Tsereteli (1977; 1978).

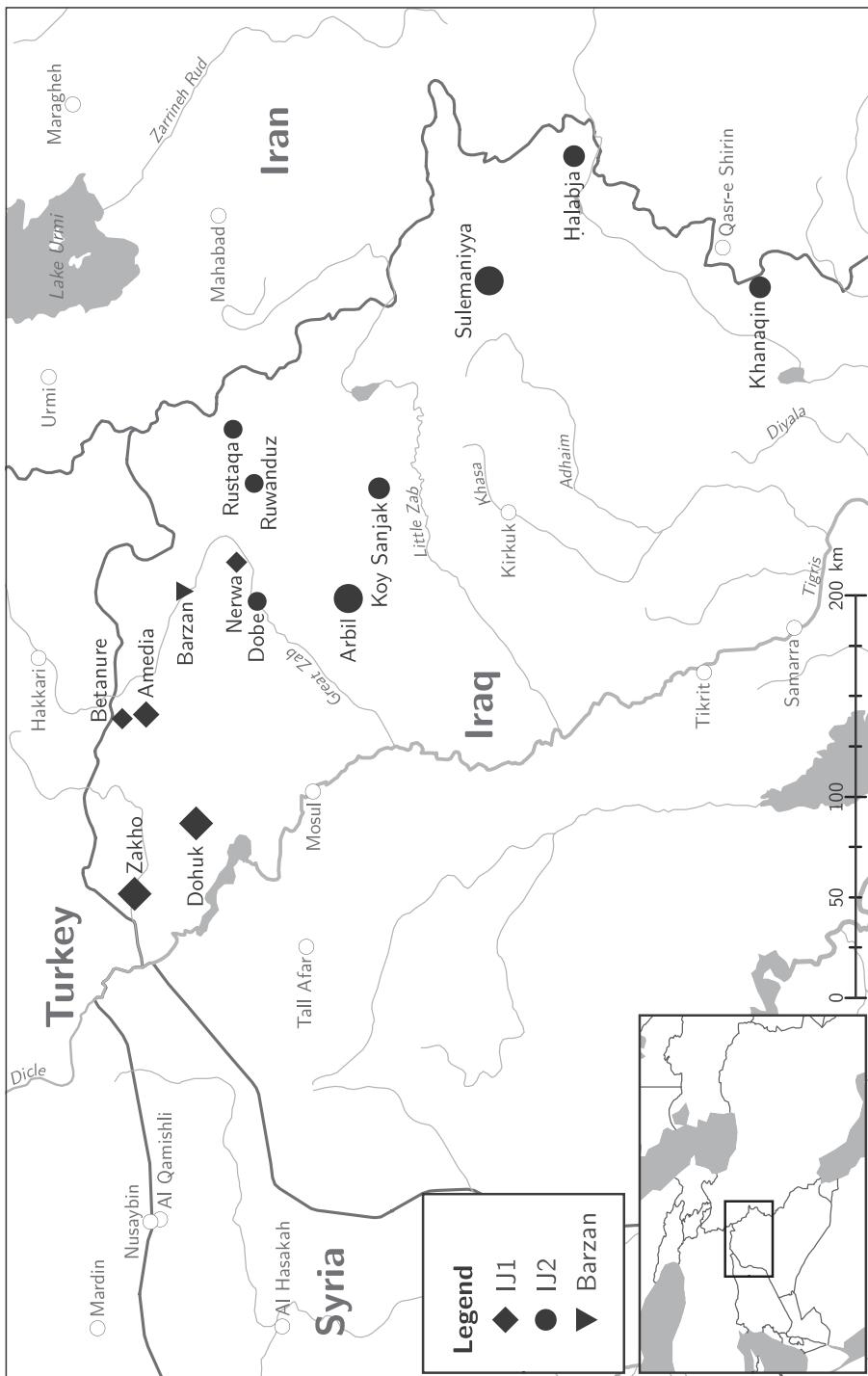


Figure 1: Locations of the Jewish NENA dialects mentioned in this chapter

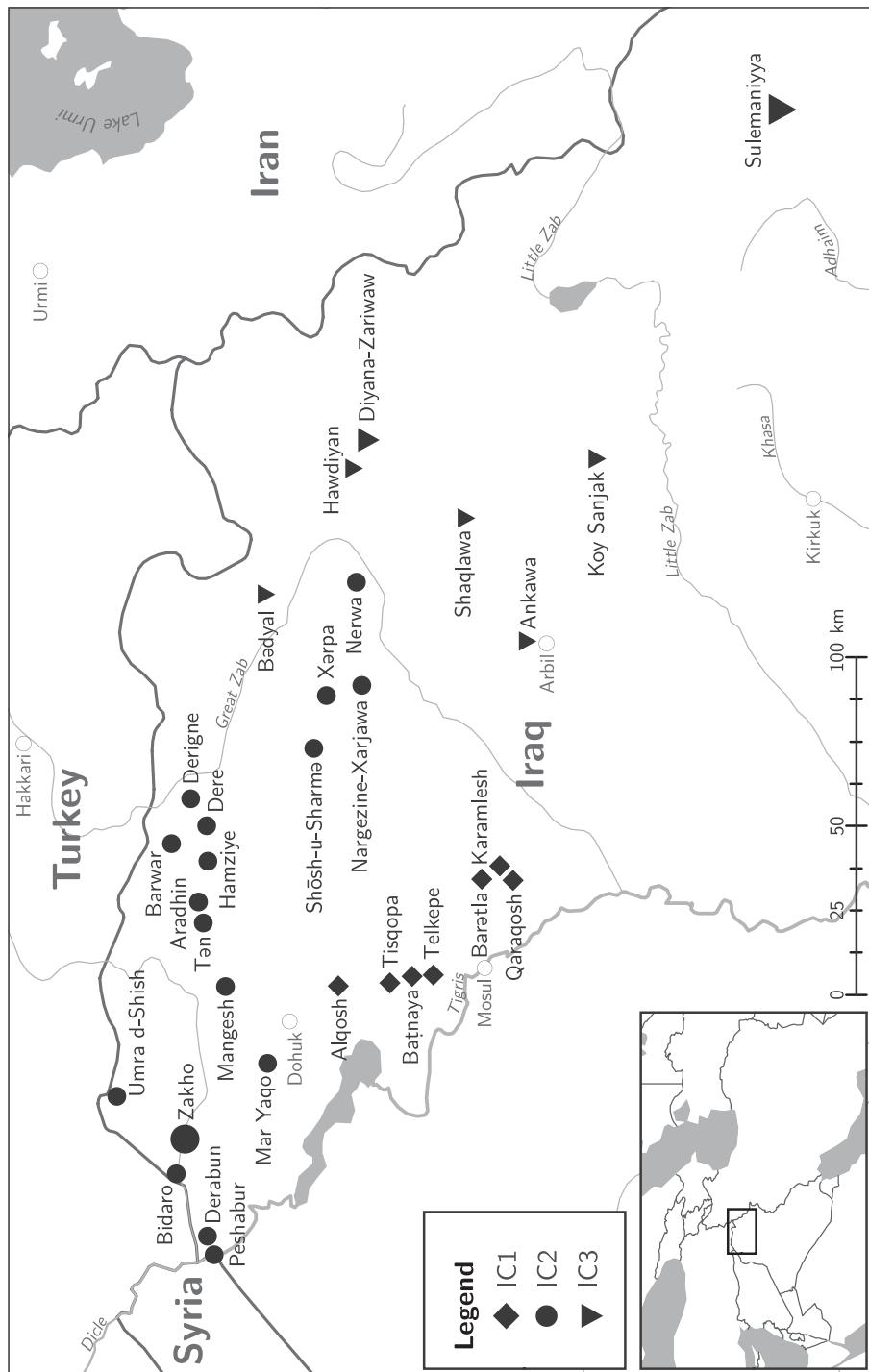


Figure 2: Locations of the Christian NENA dialects mentioned in this chapter

2. History of the speech community

The Christian and Jewish Aramaic-speaking communities in northern Iraq have deep historical roots in the region. This is reflected by the linguistic background of the NENA dialects, which, as remarked above, are shown by a number of features that originate in vernacular Aramaic in antiquity when Akkadian was still a living language. The Christians are descendants of communities that converted to Christianity in the first half of the 1st millennium C.E. (Fiey 1965). Until the end of the 1st millennium C.E. Aramaic was spoken more widely by Christians and Jews throughout Iraq. By around the 10th century the communities in the centre and south of the country had largely become Arabic-speaking. In the course of the Middle Ages many of the Christians of central Iraq migrated to the north of the country to avoid forced conversion with its consequent arabicization. We know, for example, that around the eleventh century a large proportion of the Christian population of Tikrit migrated northwards and settled in the Mosul region, many in the town of Qaraqosh (Khan 2002a: 2–3). In the Timurid invasions of the fourteenth century the Christian communities of central and southern Iraq were decimated, which would have hastened the end of any vestiges of Aramaic in this region. However, the Christian communities in the mountains in the north of the country and in villages on the nearby plain who fled to the mountains were largely spared the ravages of these invaders. There are only scant historical records concerning the Jews of northern Iraq before the modern period. Mann (1931: 216) publishes letters datable to the beginning of the sixteenth century that mention villages in the highlands of northern Iraq that had Jewish communities. Some of these had been abandoned by their Jewish population by the twentieth century or the number of Jewish inhabitants was considerably reduced, reflecting the fact that the Jewish population of the region declined after the sixteenth century. One reason for this reduction is likely to have been the forcible conversion of some Jews to Islam, especially in the nineteenth century in some areas (Soane 1912: 186).

The momentous events of the twentieth century brought about the destruction and displacement of a large number of the Aramaic-speaking communities of the region. The decisive event for the Christians was the First World War, during which the Christians in the mountains of what is now southeastern Turkey and the Urmi region in northwestern Iran sided with Russia against the Ottoman government. As a result, their villages were destroyed by the Turkish army and they were either massacred or driven from their homes (Gaunt 2006; Rockwell 1916). A large proportion of the survivors settled in refugee camps, villages and towns in Iraq. Some of the Aramaic-speaking communities in northern Iraq are, therefore, relatively recent immigrants from southeastern Turkey and northwestern Iran. The foundation of the State of Israel in 1948 was the decisive event for the Aramaic-speaking Jews of the region, since shortly thereafter the entire Jewish population of Iraq emigrated to the new state, except for a few isolated individuals, most of whom

converted to Islam. In recent decades in Iraq there has been further displacement of the Aramaic-speaking Christian communities. In the late 1970s and early 1980s, for example, many of the numerous Christian villages in northern Iraq and south-eastern Turkey were destroyed by the Iraqi and Turkish armies during the Kurdish uprisings. Further displacements of Christian communities from the villages of the Mosul plain have taken place recently. These recent events have driven many of the Christian communities to leave Iraq and settle elsewhere, mainly in Europe, North America and Australia.

3. Current status of the speech community

As remarked, the NENA subgroup is extremely diverse and the size of the speech community of individual dialects varies considerably. Some dialects are now reduced to a handful of final speakers. This applies in particular to the Jewish dialects, all of which are now highly endangered and will not survive much beyond the next two decades. Several Jewish dialects have recently become extinct, e. g. Nerwa in 2012, Sandu in 2010, Challa in 2007, Shahe in 2000, Bədyl in 1998 (Mutzafi 2014). Some dialects of small Christian communities in Iraq are also highly endangered, e. g. the Christian dialect of Bədyl. In general, however, Christian dialects are generally less endangered than Jewish ones. It is not possible to give precise statistics for individual dialects, but some of the larger ones have several thousand speakers, taking into account speakers in migrant communities. Among Christian Aramaic-speakers from Iraq a particularly widely spoken dialect is a koine that developed in the towns after the merging of various refugee communities after the First World War (Odisho 1988: 19–38).

The Aramaic-speaking Christians of Iraq belong to various church denominations, including the Assyrian Church of the East, Ancient Church of the East, Chaldean, Syriac Orthodox and Syriac Catholic.

Before the disturbances of the First World War the majority of Christians were agriculturalists living in villages. After the settlement of refugees in the towns many adopted various urban professions. Most of the Aramaic-speaking Jews were town-dwellers, who were small traders, goldsmiths, tailors, weavers and dyers. Some of the traders were shopkeepers while others were pedlars who hawked their wares around the surrounding countryside. Some of the Jews who remained in villages down to the twentieth century, such as the communities of the villages of Betanure, Shukho and Sandu, were agriculturalists (Mutzafi 2008a, 2014).

4. Sociolinguistic situation

All speakers of Aramaic who were resident or originated in northern Iraq are bilingual and in many cases trilingual. The Christians of the Mosul plain region speak Arabic in addition to their local Aramaic dialects. They have been exposed to both local Arabic vernaculars and to the literary Arabic of the Iraqi education system. In many cases Arabic has had a major impact on their Aramaic speech, particularly in peripheral dialects such as Qaraqosh, where code-switching between Aramaic and Arabic is common. Aramaic-speakers elsewhere are generally fluent in Kurdish, the language of the majority of the surrounding Muslim population, and those who have been through the Iraqi education system know Arabic. In some areas Aramaic-speakers spoke Turkoman, especially Turkoman converts to Christianity.

Although the spoken NENA dialects have their roots in antiquity, they were not committed to writing in any systematic way until the seventeenth century, from which period onwards there are extant texts written by Christians (Mengozzi 2002; Mengozzi and Braida 2011) and Jews (Sabar 1976).

The literary form of Christian NENA that emerged in the seventeenth century was used by Chaldean Christians in the region of Mosul down to modern times (Macuch 1976). Another type of Christian literary NENA emerged in the nineteenth century in Urmi, in northwestern Iran. This was originally developed by western missionaries to that region in order to propagate Bible translations in the vernacular, but subsequently gave rise to a very variegated literature (Murre-van den Berg 1999). This form of written NENA was based on the spoken Christian dialect of the Urmi region, but came to be used widely by Christians in Iraq. Within the last few decades there has been an increase in the production of literature written in this form of Christian literary NENA. Furthermore education through the medium of the literary language has, in recent years, undergone considerable development in northern Iraq. All forms of literary NENA written by Christians are influenced by the classical Syriac language, in its orthography, grammatical structure and lexicon.

The written form of the Jewish NENA dialects, by contrast, is not influenced by a classical form of Jewish Aramaic. It has continued to be produced down to modern times, especially to commit to writing traditions of oral literature (Aloni 2014). No systematic form of education system in Jewish NENA developed in the communities. Within Israel, in recent decades, however, the communities have organized various cultural events in an attempt to keep their spoken dialects alive. These include theatre plays, poetry readings and other types of entertainment.

5. Phonology

The phonology of the NENA dialects of northern Iraq exhibit some historical developments that are common to the majority of the NENA dialects and some that are particular to individual dialects. In many cases such developments have been stimulated by language contact. Dialects differ in the degrees of conservatism in the phonology, the less conservative dialects eliminating a greater degree of sounds that do not occur in contact languages. In what follows a selection of cases will be presented from the consonant system. The term ‘earlier Aramaic’ is used to refer to what can be assumed to be the ultimate historical form of features attested in NENA.

5.1. The consonant group *bgdkpt*

In earlier Aramaic the stop consonants *bgdkpt* developed fricative allophones after vowels, which can be represented *[b], *[g], *[d], *[k], *[p], *[t]. In the NENA dialects these fricative allophones became phonemicized, with the result that minimal pairs are found with stops and fricatives, e. g. C. Qaraqosh (IC1):⁵ *šata* ‘year’—*šaθa* ‘fever’; *guda* ‘wall’—*guða* ‘churn’. In the NENA dialects there is a general uniformity of the reflexes of the historical fricatives *[b], *[p] and *[k]. These reflexes are as follows:

- | | |
|------------|---|
| *[b] > /w/ | C. Qaraqosh (IC1) <i>sawa</i> (< *sābā) ‘grandfather’ |
| *[p] > /p/ | C. Qaraqosh (IC1) <i>kipa</i> (< *kēpā) ‘stone’ |
| *[k] > /x/ | C. Qaraqosh (IC1) <i>baxə</i> (< *bākē) ‘he weeps’ |

In some dialects in southeastern Turkey /x/ has subsequently shifted to the pharyngeal /ħ/ (Khan, this volume, chapter 2.5, §5.1).

A shared innovation of NENA that distinguishes the subgroup from the neighbouring Turoyo subgroup of Neo-Aramaic is the loss of the voiced velar fricative [g]. In the NENA dialects of Iraq the reflexes of this are generally either the laryngeal stop // or zero. The historical development was *[g] > // > /' / > Ø,⁶ in which the velar first shifted to a pharyngeal fricative before weakening to a laryngeal and

⁵ The abbreviations C. and J. are used to denote Christian and Jewish dialects respectively. Unless otherwise indicated, the data in this chapter are taken from the published descriptions of NENA dialects. These include Khan (2002a C. Qaraqosh), Khan (2008 C. Barwar), Khan (1999 J. Arbel), Khan (2004 J. Sulemaniyya), Talay (2001 C. Nerwa), Coghill (2003 C. Alqosh), Greenblatt (2011 J. Amedia), Mutzafi (2004a J. Koy Sanjak), Mutzafi (2008a J. Betanure), Napiorkowska (2015b C. Diyana-Zariwaw). Some material is taken from the data of the NENA database project gathered by G. Khan, E. Coghill and R. Borghero.

⁶ The velar fricative *[g] was an allophone of /g/ in earlier periods of Aramaic, and so is written in square brackets. The modern reflexes // and /' /, by contrast, are phonemes.

zero (Tsereteli 1990). The reflex /'/ is, accordingly, more archaic than the reflex Ø. The distribution of /'/ differs across the Iraqi NENA dialects:

- C. Qaraqosh (IC1): *šra'a* (< *šrāḡā) ‘lamp’, *ra'əš* (< *rāḡeš) ‘he wakes’, *lə'ma* (< *laḡmā) ‘jaw’, *palə'* (< *pāleg) ‘he divides’, *pal'a* (< pālḡā) ‘she divides’
- C. Barwar (IC2): *šraya*, *rayəš*, *pela* (< *paḡlā) ‘radish’, *păle*, *păla*
- C. Diyana-Zariwaw (IC3): *šreya*, *rayəš*, *lama*, *păle*, *păla*
- J. Amedia (IJ1): *šra'a*, *ra'əš*, *pe'la*, *pale'*, *pal'a*
- J. Koy Sanjak (IJ2): *zo'a* (< *zōḡa) ‘pair’, *pella*, *ră'əš*, *păle*, *palya*

The pharyngeal is preserved in a few words in some Jewish dialects in pharyngealized environments, e. g. J. Amedia *l-‘oya*, J. Koy Sanjaq *lo'a* (< *lə-ḡawwāyā) ‘inside’.

A considerable diversity is exhibited by the NENA dialects of Iraq in the reflexes of the interdental fricatives *[t] and *[d]. These are preserved in most of the Christian dialects of IC1 and IC2, e. g.

- C. Qaraqosh (IC1): *xaθa* (< *'aḥāṭā) ‘sister’, *iða* (< *'ēdā) ‘festival’
- C. Barwar (IC2): *xaθa* ‘sister’, *eða* ‘festival’

In the IC2 dialect of Nerwa the unvoiced /θ/ has merged with the voiced /ð/, e. g.

- C. Nerwa (IC2): *xaða* ‘sister’, *eða* ‘festival’

Some Christian dialects in the far northwestern periphery of IC1 and IC2 in and around Zakho and adjacent to the Tigris replace the interdentals with stops or sibilants, e. g.

- C. Bidaro: *xata* ‘sister’, *'eda* ‘festival’
- C. Peshabur: *xasa* ‘sister’, *'eza* ‘festival’

Dialects on the eastern periphery of IC2 and most dialects in IC3 to the east of the Zab have replaced the interdentals with stops or sibilants, e. g.

- C. Xərpa (IC2): *xasa* ‘sister’, **'iza* (< *'īdā) ‘hand’
- C. Bədylə (IC3): *xata* ‘sister’, *'ida* ‘hand’

In some Christian dialects the unvoiced interdental is replaced by a sibilant and the voiced interdental by a stop, but never the reverse, e. g.

- C. Sulemaniyya (IC3): *xasa* ‘sister’, *'eda* ‘festival’

Among the Christian IC3 dialects there are, however, some archaizing dialects that preserve interdentals, in some cases only the unvoiced one, e. g.

- C. Ankawa (IC3): *xaθa* ‘sister’, *'ida* ‘hand’

Both interdentals are preserved in the speech of Christians in Hawdiyan (northeast Iraq) who originate from Para and Shapat in southeastern Turkey.⁷

There is some conservatism of the interdentals among the Jewish dialects of the IJ1 group, e. g.

J. Dohuk (IJ1): *xaθa* ‘sister’, *iða* ‘hand’

Most dialects of IJ1, however, exhibit some degree of replacement of interdentals with either stops or sibilants. If there is conservatism, it is of the unvoiced interdental, e. g.

J. Betanure/J. Amedia (IJ1): *xaθa* ‘sister’, *'ida* ‘hand’

J. Nerwa (IJ1): *xasa* ‘sister’, *'ida* ‘hand’

J. Zakho (IJ1): *xasa* ‘sister’, *'iza* ‘hand’

The Jewish trans-Zab IJ2 group of dialects have the distinctive feature of the shift of interdentals to the lateral /l/, e. g.

J. Arbel (IJ2): *mala* (< **mātā*) ‘village’, *'ila* (< **'idā*) ‘hand’

In some dialects the velar stops /k/ and /g/ are palatalized to the affricates /č/ [tʃ ~ ts] and /j/ [dʒ ~ dz]. This is found predominantly in the northeastern sector of the region in the IC3 group, e. g. C. Diyana-Zariyaw *čalba* (< **kalba*) ‘dog’, *jare* (< *gare*) ‘roof’, C. Bədylal *čalo* (< **kalo*) ‘bride’, *julpa* (< **gulpa*) ‘wing’. Some dialects also in IC2 exhibit some degree of palatalization in these consonants before unrounded vowels, e. g. C. Barwar *k'asa* ['c̥yæ:sæ] (< **kasa*) ‘stomach’, *g'awra* ['j̥yæwræ] (< *gawra*) ‘man’.

In the C. Bədylal dialect in the northeastern sector of the region in the IC3 group the alveolar stops *t and *d are palatalized to the affricates /č/ [tʃ] and /j/ [dʒ] before high vowels, e. g. C. Bədylal *beta* ‘house’, *beči* (< **beti*) ‘my house’, *'ida* ‘hand’, *'iji* (< **'idi*) ‘my hand’.

In dialects in the north of the region in the area where Kurmanji Kurdish is spoken, there are several Kurmanji loanwords with unaspirated unvoiced stops, e. g. C. Barwar (IC2) *tanga* ‘saddle girth’ (where t = unaspirated) < Kurmanji *teng*. In these NENA dialects some unaspirated stops have developed in NENA words, usually where a voiced stop is followed by an adjacent laryngeal stop, e. g. C. Barwar *t-ile* ‘that he is’ < *d-ile*.

5.2. Pharyngeals

The pharyngeal consonants *‘ (voiced pharyngeal fricative) and *h (unvoiced pharyngeal fricative) generally undergo the following shifts in NENA dialects:

⁷ Lidia Napiorkowska (pc.).

*^c > / / or Ø

*h > /x/

The development of the voiced pharyngeal *^c is parallel to that of *[g̣] and the distribution of the reflexes / / and Ø in the dialects of Iraq is the same as described above for the reflexes of *[g̣], e. g.

C. Qaraqosh (IC1): *zra'a* (< *zrā'ā) ‘cultivation’, *da'or* (< *dā'er) ‘he returns’, *bə'ta* (< *bē'tā) ‘egg’, *šamə* (< *šāme) ‘he hears’, *tar'a* (< tar'ā) ‘door’

C. Barwar (IC2): *zraya*, *dayər*, *bita*, *šāme*, *tāra*

C. Diyana-Zariwaw (IC3): *zraya*, *ṭayər*, *bita*, *šāme*, *tāra*

J. Amedia (IJ1): *pqa'a* (< *pqā'ā) ‘explosion’, *da'or*, *be'ia*, *šame*, *tar'a*

J. Koy Sanjak (IJ2): *be'e* (< *bē'ē) ‘eggs’, *da'or*, *beta*, *šāme*, *tāra*

The development of the unvoiced pharyngeal *h to the velar fricative /x/ is an innovation of the NENA subgroup and is exhibited by all dialects of Iraq, e. g.

C. Barwar (IC2): *xmara* (< *hmārā) ‘ass’, *xaθa* (< *'ahātā) ‘sister’

In the far northwest close to the Turkish border the secondary shift of /x/ > /h/ is found in the dialect of Derabun, e. g. *hata* (< *xaθa < *'ahātā) ‘sister’. The shift has also been applied to cases of /x/ that are reflexes of *[k], e. g. *bahe* (< *baxe < *bākē) ‘he weeps’. This is a feature of a cluster of NENA dialects in southeastern Turkey (Khan, this volume, chapter 2.5, §5.1), from where the Derabun community originated.

In some of the NENA dialects of Iraq the original voiced and unvoiced pharyngeal are preserved in words containing emphatic, i. e. pharyngealized, consonants, including /q/. This is found predominantly in the groups IC1, IC3, IJ1 and IJ2, but not in the Christian dialect group IC2, e. g.

C. Qaraqosh (IC1): *ta'en* (< *tā'en) ‘he lifts’, *rahoqa* (< *rāhōqā) ‘distant’

C. Sulemaniyya (IC3): *ta'el* (< *tā'el) ‘he plays’, *haruqa* (< *rāhōqā) ‘distant’

J. Zakho (IJ1): *'apṣa* (< *'apṣā) ‘gallnut’, *rahuqa* (< *rāhōqā) ‘distant’

J. Koy Sanjak (IJ2): *ta'na* (< *ta'nā) ‘load’, *ḥānəq* (< *ḥāneq) ‘he chokes’

This preservation has come about by a process of ‘long distance consonant agreement’ (Khan 2013). As can be seen above, in J. Koy Sanjak (IJ2), the preservation of the pharyngeal / / entailed the neutralization of the pharyngealization of the emphatic *t resulting in plain /t/ (Mutzafi 2004a: 28–33) and in C. Sulemaniyya the pharyngealization is lost from the *t resulting in an unaspirated plain /t/.

5.3. Pharyngealized consonants

In some of the NENA dialects of Iraq the emphatic pharyngealized consonants **t* and **s* of earlier Aramaic have been preserved as pharyngealized segments. This is found predominantly in dialects in groups IC1, IC2 and IJ1 in the western areas for the region, e. g.

C. Qaraqosh (IC1): *tina* (< **ṭinā*) ‘mud’, *tarpa* (< **ṭarpa*) ‘leaf’, *səpra* (< **ṣeprā*) ‘sparrow’, *sliwa* (< **ṣlibā*) ‘cross’

The pharyngealization of the emphatic consonants in such dialects tends to spread to adjacent segments in ways that are similar to what has been documented in Arabic dialects. As in some Arabic dialects (Hoberman 1989; Watson 1999), the spreading tends to be restricted to segments in adjacent syllables and is blocked by high front vowels (Khan 2013).

In dialects in groups IC3 and IJ2, which were spoken in northeastern areas of the region, the pharyngealization of the historically emphatic consonants **t* and **s* tends to be weakened.

In the Jewish dialects of IJ2 /*t*/ and /*s*/ are weakly pharyngealized or lose their pharyngealization altogether. The pharyngealization is neutralized in particular when the word contains a pharyngeal consonant, e. g. J. Koy Sanjak (IJ2) *ta'na* (< **ta'na*) ‘load’. This process of development seems to be due to convergence with the behaviour of pharyngealization in the Kurdish dialects described by Margaret Kahn (1976: 49–52). In Kahn’s analysis a pharyngealized consonant in Kurdish is the result of the presence of an underlying pharyngeal segment in the word. Such words can also be realized with the pharyngeal on the phonetic surface in place of the pharyngealized consonant. Kahn found this type of pharyngealization in Kurmanji Kurdish dialects north of Urmi. It is significant, however, that today it does not seem to exist in the Sorani dialects of Iraq or Iran. The feature is found in the Jewish NENA dialects that were in Sorani Kurdish areas (i. e. 1J2 group in Iraq and the Jewish dialects of western Iran, see Khan, this volume, chapter 4.4). It is possible, therefore, that this typology existed at an earlier historical period in Sorani Kurdish and the NENA dialects converged with this typology and preserved it whereas it was lost in Sorani Kurdish.

In many of the Christian dialects in the northeastern region of Iraq (IC3) the pharyngealization of the historical emphatic consonants **t* and **s* is totally lost in some words, e. g. C. Diyana-Zariyaw (Napiorkowska 2015a) *seda* (< **ṣaydā*) ‘hunt’. Reflections of the presence of the original emphatic phonemes have, however, been preserved in a number of features. These include the lack of aspiration of the original **t* after loss of emphasis, e. g. *xatte* (< **haṭṭē*) ‘wheat’, and backing and/or lowering of adjacent vowels, e. g. *xpsa* (< **ḥāṣā*) ‘back’, *sɔma* [sɔ:ma] (< **ṣawmā*) ‘fast (noun)’.

Pharyngealization is also associated with a group of non-historical emphatic

consonants that includes the sonorants /r/, /l/ and the labials /b/, /m/, /p/, also /w/. This group developed in certain words, in some cases, it seems, to fulfil the functional need of creating semantic contrasts, e. g.

- (1) J. Amedia (IJ1)

<i>mlele</i>	'he filled'	<i>mlele</i>	'it sufficed'
<i>baza</i>	'hawk'	<i>b-aza</i>	'she will go'
<i>bärke</i>	'pool'	<i>bärke</i>	'his knee'

The phonemic contrast between the plain and emphatic rhotics /r/ : /ṛ/ is particularly common in the NENA dialects. In a few cases the number of rhotic phonemes has proliferated to three. In Iraq this is found, for example, in the dialect of C. Barəṭla (IC1), which has an emphatic /ṛ/, a plain tap /r/ and a retroflex /ṛ/ (Arnold 2002; Mole 2015):

<i>pare</i> < * <i>pārā</i> 'bran, husks left after sieving'
<i>pare</i> < * <i>parre</i> 'lambs'
<i>pare</i> < * <i>pāre</i> 'money' (< Kurd. <i>pāra</i>)

5.4. Vowels

Various shifts have taken place in the quality of vowels across the NENA dialects of Iraq. These include the following.

*/ē/

An original *ē has shifted to /i/ in word-medial position in a few dialects. These include C. Qaraqosh (IQ1) on the southwestern periphery and some Christian dialects in the northeastern sector in group IC3, e. g. C. Qaraqosh (IC1), C. Diyana-Zariwaw (IC3) *riša* (< **rēšā*) 'head'.

In C. Qaraqosh */ē/ in a final open unstressed syllable undergoes centralization to /ə/, e. g. *xmarə* (< **hmārē*) 'asses'.

*/e/

The reflex of original short */e/ is generally a centralized /ə/, e. g. C. Qaraqosh (IC1) *šaqəl* (< šāqel) 'he takes'. In many dialects /ə/ is also the reflex of a shortened */i/, e. g. C. Qaraqosh *xləmta* (< **hlīmtā*) 'thick', but C. Barwar *xlimta*.

*/u/, */ū/

In several dialects in the northeastern sector of the region there is a tendency to front */u/ and */ū/ to the region of [y] or [u], represented here as /ü/, e. g. C. Diyana-Zariwaw *ürxa* (< 'urhā) 'road', *bariūza* (< **barrūzā*) 'dry'.

*/ð/

In some Christian dialects in the northeast of the region (IC3) *ð is raised to /u/, e. g. C. Diyana-Zariyaw (IC3) *muxa* (< *mōħā) ‘brain’, C. Bədyl (IC3) *şadure* (< *şadore) ‘to send’. In such dialects the vowel sometimes undergoes both raising and fronting, e. g. C. Bədyl *xüna* (< *'ahōnā) ‘brother’.

*/a/ and */ā/

The distinction between the historical short */a/ and long */ā/ phonemes has been lost. These collapsed into a single phoneme /a/, the length of which is determined by syllable structure. In some dialects this /a/ phoneme underwent qualitative changes in certain contexts.

In some dialects, predominantly in the northern sector of the region, /a/ is raised to /e/ before the laryngeals /ʃ/ (< */ʃ/) and /h/, e. g. J. Amedia (IJ1) *te'na* (< *ta'nā) ‘load’, *sehra* (< *sahrā) ‘moon’. This has occurred in dialects that subsequently lost the laryngeals, e. g. C. Barwar (IC2) *tena* ‘load’, *sera* ‘moon’, C. Diayana-Zariyaw (IC3) *zdetə* (< *zda'tā) ‘fear’. In some of these dialects there is raising of /a/ also before an originally geminated consonant in a few lexemes, e. g. C. Barwar (IC2) *genawa* (< *gannābā) ‘thief’, C. Diayana-Zariyaw *ginawa*; cf. C. Qaraqosh (IC1) and J. Amedia (IJ1) *ganawa*.

An /a/ in final unstressed open syllables undergoes backing in C. Telkepe (IC1), e. g. *tamp* ['tæ:mɒ] (< *tāmā) ‘there’ (Coghill 2008).

*/aw/

The diphthong */aw/ contracts in most dialects to /o/, e. g. C. Qaraqosh (IC1) *moθa* (< *mawtā) ‘death’. In dialects in the northeastern sector of the region this /o/ may undergo fronting, e. g. C. Diyana-Zariyaw *jöza* (< gawzā) or raising, e. g. C. Bədyl *muta* ‘death’. In some dialects in the northern sector of the region this diphthong has been preserved, e. g. C. Barwar *mawθa* ‘death’.

*/ay/

The diphthong */ay/ contracts in most dialects to either /e/ or /ɛ/, e. g. C. Qaraqosh (IC1) *beθa* (< *baytā) ‘house’, C. Barwar *beθa*. Some dialects in the far northwest do not monophthongize the diphthong completely, e. g. C. Bidaro *beyta* ‘house’.

5.5. Syllable structure

The onset of syllables may consist of a single consonant *C* or a cluster *CC* (with certain constraints, see below), e. g. (syllable boundaries are marked by a dot):

- (2) C. Barwar (IC2)
 - be.θa* ‘house’
 - smo.qa* ‘red’

A syllable in principle does not have a vowel as its onset but always begins with at least the laryngeal //, e. g. C. Barwar *'á.na* ‘I’, *'á.ti* ‘you’, *'i.lána* ‘tree’. In some trans-Zab Jewish dialects (IJ2) an initial // shifts to /h/, e. g. J. Sulemaniyya *hít* ‘there is’ (< *'it*), *heka* ‘where?’ (< *'eka*).

There are differences across the dialects of Iraq with regard to the presence of the laryngeal glottal stop // in clusters in the syllable onset. Some dialects, predominantly in groups IC1 and IJ1, tolerate // as the first or second segment of the initial cluster, e. g.

C. Qaraqosh (IC1): *'si.ra* ‘tied’, *r'i.ša* ‘awake’

J. Amedia (IJ1): *'wi.da* ‘done’, *r'i.ša* ‘awake’

Dialects of other groups tend not to tolerate // in initial clusters and elide them in these circumstances.

The cluster is sometimes pronounced without being broken by an epenthetic vowel. This is generally the case if the first or second segment is a sibilant or sonorant continuant, e. g. C. Barwar *šte.li* ‘I drank’, *pli.te.le* ‘he has gone out’. Otherwise clusters tend to be broken by an ultrashort epenthetic vowel, e. g. C. Barwar *kpi.na* [k³p̥i:na] ‘hungry’. An epenthetic prosthetic vowel is sometimes pronounced before a cluster, especially if the first segment is a sonorant or sibilant, e. g. C. Qaraqosh *nqa.ša* [ənqa:ʃa] ‘beating’, *ski.na* [əski:na] ‘living’. An alternative strategy of treating a cluster of two stops is to elide the first of them. This is found in C. Bədyl, e. g. *pina* (< **kpinā*) ‘hungry’, *tāna* (< **ptāna*) ‘plough’.

Across the NENA dialects several monosyllabic nouns have been made bisyllabic by geminating a consonant after an epenthetic vowel breaking an initial consonantal cluster, e. g.

(3) C. Barwar

šám.ma ‘name’ < **šmā*

dám.ma ‘blood’ < **dmā*

Monosyllabicity is tolerated to a greater extent in other grammatical categories.

Historical gemination of consonants has been lost after the low vowel /a/ but is generally preserved after high vowels, especially /ə/, e. g.

(4) C. Barwar

kaka ‘tooth’ < **kakkā*

'əzza ‘goat’ < **'əzzā*

The sequence *VCCCV* in word internal position is found in certain inflections of verbs. It may be syllabified *VCC.CV* or *VC.CCV*. An epenthetic vowel is usually inserted between the *CC* cluster in the coda of *VCC* or the onset of *CCV*, e. g.

(5) C. Barwar

mapál.xa ~ *map.lóxa* ‘she uses’ < *maplxa*

The rhymes of syllables are in principle bimoraic with a VV or VC structure. As a result, vowels in open syllables are generally long and those in closed syllables are short, e. g. C. Barwar *'i.ði* [?i:ði:] ‘my hand’, *qim.li* [qimli:]. In the transcription, therefore, in this and other chapters on Neo-Aramaic the default length of vowels is left unmarked by diacritics, i. e. a vowel without a diacritic in an open syllable is long and a vowel without a diacritic in a closed syllable is short. The only exception is the vowel /ə/, which is always short. Diacritics are used only when the length of the vowel deviates from this principle. The surface phonetic duration of vowels is conditioned by a number of other factors, such as stress and speed of utterance. Some cases of phonetically short vowels in open syllables arise from the presence of an underlying glottal. This is seen clearly in a conservative dialect such as C. Qaraqosh (IC1), in which the glottal stop // alternates with zero, e. g.

(6) C. Qaraqosh

da'.ra [da?ra:] ~ *dă.ra* [dara:] ‘she returns’

tar.'a [t^har?a:] ~ *tăr:a* [t^hara:] ‘door’

In the alternants in which the glottal stop has been dropped the vowel length remains unchanged, which can be interpreted as reflecting the continued presence of the glottal at an underlying level.

5.6. Stress

The basic position of word stress in most NENA dialects in Iraq is on the penultimate syllable, e. g. C. Qaraqosh (IC1) *béθa* ‘house’, *k-šáqəl* ‘he takes’. Clitics, such as the copula, are not taken into account in stress placement, e. g.

(7) C. Qaraqosh

béθan=ilə

house.our=COP.3MS

‘it is our house’

Stress is generally not placed on particles affixed to verbal forms, which can result in the placement of the stress before the penultimate syllable of the word. This lack of stressing of affixes is a property of clitics. It is not the case, however, that the stress is completely unaffected by the affixes, since it moves to the end of the verbal stem. I shall continue, therefore, to refer to these elements as verbal affixes rather than clitics, e. g.

- (8) C. Qaraqosh
k-šaqél-wa-lhə
IND-take.PRS.3MS-PST-L.3PL
‘he used to take them’

In the trans-Zab Jewish dialects of the IJ3 group the basic position of stress is on the final syllable of the word, e. g. J. Sulemaniyya *belá* ‘house’. This is likely to be due to convergence with Kurdish. As in other NENA dialects, stress is not in principle placed on clitics or verbal affixes, e. g.

- (9) J. Sulemaniyya
a. *laxxá=ya*
here=COP.3FS
‘she is here’
b. *gárás-wa-lu*
pull.PRS.3MS-PST-L.3PL
‘he used to pull them’

6. Nominal morphology

6.1. Pronouns

The NENA dialects of Iraq have various series of third person pronoun. A distinction may be made between deictic pronouns, which point to a referent, and anaphoric pronouns, which signal that a referent is identifiable in the context (typically the discourse history) without pointing to its location. A syntagmatic distinction may be made between independent pronouns and adnominal pronouns that are combined with a noun. Deictic pronouns typically distinguish between near deixis and far deixis. With regard to gender and number, third person pronouns distinguish masculine singular, feminine singular and common plural.

There are differences across the NENA dialects of Iraq with regard to the degree to which these various types of third person pronouns are expressed by separate morphological forms. Some dialects are lacking distinct forms for some of the slots, as can be seen in the table below, which presents the independent third person pronouns from a selection of dialects:

Table 1: Independent third person pronouns in Iraqi dialects of NENA

	C. Qaraqosh (IC1)	C. Barwar (IC2)	C. Diyana-Zariwaw (IC3)	J. Amedia (IJ1)	J. Arbel (IJ2)
Near deixis					
3MS	' <i>aða</i>	' <i>awwa</i>	' <i>awwa</i>	' <i>ayya</i>	' <i>iyya</i>
3FS	' <i>aði</i>	' <i>ayya</i>	' <i>ayya</i>	' <i>ayya</i>	' <i>iyya</i>
3PL	' <i>anə</i>	' <i>anna</i>	' <i>anne</i>	' <i>anna</i>	' <i>anne</i>
Far deixis					
3MS	' <i>awa</i>	' <i>āwaha</i>	' <i>awó</i>	' <i>awāħa</i>	' <i>o</i>
3FS	' <i>aya</i>	' <i>āyaha</i>	' <i>ayó</i>	' <i>ayāħa</i>	' <i>o</i>
3PL	' <i>ane</i>	' <i>ānaha</i>	' <i>anó</i>	' <i>anāħa</i>	' <i>anne</i>
Anaphoric					
3MS	' <i>ahu</i>	' <i>aw</i>	' <i>o</i>	' <i>o</i>	' <i>o</i>
3FS	' <i>ahi</i>	' <i>ay</i>	' <i>e</i>	' <i>e</i>	' <i>o</i>
3PL	' <i>anhən</i>	' <i>ani</i>	' <i>an</i>	' <i>an</i>	' <i>oni</i>

In this sample of dialects, J. Arbel (IJ2) has a much smaller inventory of morphological forms than the other dialects. It lacks gender distinctions in the singular forms and also lacks a formal distinction between far deixis and anaphoric forms in the singular.

The C. Qaraqosh dialect preserves archaic forms of the singular near deixis pronouns with an interdental ('*aða*, '*aði*) and so exhibits heterogeneity in the morphological bases of the singular near deixis and far deixis forms. The other dialects in the table, by contrast, have lost these archaic forms by forming the near deixis pronoun base on the analogy of those of far deixis and anaphoric pronouns.

Some dialects have developed further morphological distinctions to express various forms of ‘intensive’ deixis, such as ‘very far deixis’ or ‘contrastive deixis’, e.g.

- (10) a. C. Barwar
 - 3MS. far '*āwáħa*, very far '*āwāħa*
- b. C. Peshabur (Coghill 2008: 97–98)
 - 3MS far '*āwáħa*, very far '*āwáħa*, extremely far '*āwāħa*
- c. C. Diyana-Zariwaw
 - 3MS far '*awó*, contrastive '*awóħa*, '*awóħani*

The anaphoric pronouns must in principle always be bound to a referent in the discourse, and so have not developed into definite articles. The Jewish trans-Zab dialects (IJ2) have borrowed a definite article suffix from Sorani Kurdish with the form *-āke*, e.g. J. Arbel *bela* ‘house’, *belāke* ‘the house’. The form *-āke* corre-

sponds to the oblique form of the Kurdish definite article (*-ākay*) rather than the nominative form (*-āka*).

The first person singular independent pronoun is of common gender and has the form *'ana* across all dialects of Iraq.

Differentiation of gender is expressed by distinct forms of the second person singular independent pronoun in many dialects in the northwestern sector of Iraq in the IC1 and IJ1 groups:

IC1: 2ms. *'ahət*, 2fs. *'ahat* (C. Qaraqosh); 2ms. *'ayət*, 2fs. *'ayat* (e. g. C. Alqosh, C. Telkepe, C. Batnaya, C. Tisqopa, C. Karamlesh)

IJ1: 2ms. *'ahət*, 2fs. *'ahat* (J. Betanure, J. Dohuk, J. Nerwa, J. Zakho)

Similar forms are found sporadically in the Christian dialects of northeastern Iraq, e. g.

IC3: 2ms. *'ayət*, 2fs. *'ayat* (e. g. C. Ankawa, C. Sulemaniyya)

In other dialects a second person singular form of common gender is used, the most widespread forms being *'at* and *'ati*.

There are no gender distinctions in the independent first and second person plural pronouns, just as there is no gender distinction in the 3rd person plural demonstrative pronouns. The most widespread forms of the first person plural independent pronoun are *'axni* and *'axnan*. Some dialects have innovative forms based on the analogy of the 2pl. pronoun, e. g. *'atxan* on the analogy of 2pl. *'atxun* (J. Arbel, J. Ruwanduz, J. Rustaqa). The most widespread form of the 2pl. pronoun is *'axtun*, with various analogical innovations attested in individual dialects, e. g.

IC1 *'axnutən* (C. Baṭnaya)

IC2 *'axtum* (C. Nargezine-Xarjawa, C. Shōsh-u-Sharmə), *'axtaxum* (C. Nargezine-Xarjawa), *'axtoxum* (C. Shōsh-u-Sharmə, C. Xərpa), *'axtutən* (C. Tən)

IC3 *'axtun* (C. Shaqlawa), *'axnoxən* (C. Sulemaniyya)

IJ1 *'axtoxun* (J. Betanure, J. Dohuk, J. Zakho)

IJ2 *'atxun* (J. Arbel, J. Ruwanduz, J. Rustaqa, J. Ruwanduz); *'axnəxun* (J. Koy Sanjak), *'axnăxun* (J. Sulemaniyya); *'atoxun* (J. Dobe)

6.2. Pronominal suffixes

Most dialects have a single set of pronominal possessive suffixes for singular and plural nouns, although historically the suffixes had different forms according to whether the noun stem was singular or plural. Some dialects in IC1, however, have preserved a formal distinction in part of the paradigms. One such dialect is C. Qaraqosh, which has two forms of suffixes for the plural persons:

Table 2: Pronominal possessive suffixes in C. Qaraqosh

	On singular nouns	On plural nouns
3MS	-əħ	-əħ
3FS	-ah	-ah
3PL	-ħən	-ehən
2MS	-ux	-ux
2FS	-ax	-ax
2PL	-xun	-exun
1S	-i	-i
1PL	-an	-enan

Many dialects that have only one set of suffixes generalize the use of forms that were originally attached to plural nouns. This applies especially to the 3pl. and 2pl., in that they use forms related to *-ehən* and *-exun* in the paradigm above, e. g.

- 3pl. -eyħən (C. Mar-Yaqo), -ehən (C. Baṭnaya, C. Mangesh, C. Umra d-Shish), -ehən (C. Aradhin), -eyħən (C. Peshabur)
 2pl. -exu (C. Barwar), -exun (J. Nerwa), -exün (C. Bədyl)

Variation across the NENA dialects of Iraq is found in particular in the 3rd person suffixes. Examples are given below of a selection of 3ms. forms:

- 3ms. -eh (C. Baṭnaya, C. Ankawa)
 -əħ (C. Qaraqosh, C. Alqosh, C. Karamlesh, C. Mangesh, C. Peshabur)
 -e (C. Telkepe, C. Aradhin, C. Barwar, C. Dere, C. Hamziye, C. Bədyl, J. Betanure, J. Nerwa, J. Zakho)
 -eu (J. Arbel, J. Sulemaniyya)
 -u (C. Umra d-Shish, C. Zariwaw)

In C. Sulemaniyya the gender distinction in the 3s has been lost and the original 3ms form *-e* now is used with common gender reference. This matches the collapse of gender distinction in the demonstrative pronouns in this dialect (e. g. '*o gora* 'that man', '*o baxta* 'that woman').

6.3. Inflection of nouns and adjectives

In all NENA dialects nouns are either masculine or feminine in gender. Most masculine nouns have the inflectional ending *-a*, e. g. C. Barwar *sus-a* 'horse'. This, however, is not a gender marker, but rather a nominal inflection of unspecified gender, since a certain group of feminine nouns also end in *-a*, e. g. C. Barwar *'en-a* 'eye'. A large proportion of feminine nouns have marked feminine inflectional endings with one of the variant forms *-ta* and *-θa*, or phonetic developments of these, e. g. C. Barwar *sus-ta* 'mare', *slo-θa* 'prayer', *šən-θa* 'sleep', J.

Arbel *slo-la* ‘synagogue’ (< **sloθa*), *šən-da* (< **šənθa*). Some loanwords are left without being morphologically adapted and do not have these inflectional endings. All loanwords, however, are assigned a gender for agreement purposes, even if they are from languages that do not distinguish gender, such as Sorani Kurdish (Khan 2004: 177–180, 2009: 180–184).

The plural of nouns is marked by a variety of inflectional endings, which replace the aforementioned singular endings. The most common ending is *-e* (or phonetic variants thereof), which is the unmarked plural marker, as shown by the fact that loanwords that are inflected with a plural marker are generally given this ending. It is used on many masculine nouns with singulars in *-a*, on some feminine nouns ending in *-a* or *-ta* and on loanwords ending in zero in the singular, e. g. C. Barwar *kep-a* (masc.) ‘stone’, *kep-e* ‘stones’, *'ilan-a* (fem.), pl. *'ilan-e* ‘trees’, *nun-ta* (fem.) ‘fish’, *nun-e* ‘fish (pl.)’, *čangəl* (fem. sing.) ‘fork’ (< Kurd. *çengel*), *čangal-e* ‘forks’. There is no clear correlation, therefore, between the unmarked plural ending *-e* and the gender of the noun.

Nouns ending in *-a* in the singular may form plurals by various other endings, which may be categorized as marked plurals, e. g.

(11) C. Barwar

<i>'umra</i> (masc.) ‘church’	<i>'umrane</i> ‘churches’
<i>'aqla</i> (fem.) ‘leg’	<i>'aqlaθa</i> ‘legs’
<i>xona</i> (masc.) ‘brother’	<i>xonāwaθa</i> ‘brothers’
<i>beθa</i> (masc.) ‘house’	<i>beθwaθa</i> ‘houses’

Feminine nouns ending in *-ta* (or variants thereof) form plurals by an assortment of plural endings, e. g.

(12) C. Barwar

<i>karta</i> ‘load’	<i>karaθa</i> ‘loads’
<i>qalθa</i> ‘basket’	<i>qalaθa</i> ‘baskets’
<i>susta</i> ‘mare’	<i>susyaθa</i> ‘mares’

As can be seen, there is some correlation between these marked plural endings and the gender of the noun. Some of the marked plural endings are associated, moreover, with particular semantic classes of nouns. The ending *-āwaθa*, for example, in C. Barwar is commonly used with kinship terms, and the ending *-ane* is used predominantly with inanimate nouns.

In dialects on the Mosul plain (IC1) the Arabic plural ending *-at* is sometimes added to feminine nouns with singulars ending in *-a*. Most, but not all, such nouns are loanwords, e. g.

- (13) C. Qaraqosh
- | | |
|---------------------------|------------------------|
| <i>fárda</i> (f.) ‘load’ | <i>fárdat</i> ‘loads’ |
| <i>hénna</i> (f.) ‘thing’ | <i>hánnat</i> ‘things’ |
| <i>'ilána</i> (f.) ‘tree’ | <i>'ilánat</i> ‘trees’ |

Some nouns, mostly those with the bisyllabic pattern *CVCCa* in the singular, form plurals by reduplicating the final consonant, e. g. C. Barwar *təlpa* ‘eyelash’, *təlpape* (eyelashes).

Adjectives are inflected for gender in the singular and inflect for the plural by the unmarked plural ending *-e*. Unlike nouns, adjectives do not have marked plural endings e. g.

- (14) C. Barwar
- | |
|---------------------------------|
| <i>basim-a</i> (ms.) ‘pleasant’ |
| <i>basim-ta</i> (fs.) |
| <i>basim-e</i> (pl.) |

6.4. Nominal annexation

A noun may be modified by another noun in an annexation construction. In most dialects this is formed by the general subordinating particle *d*, which has been inherited from earlier Aramaic. Some dialects preserve the original syntax of the particle and sometimes prefix it to the dependent noun, e. g.

C. Qaraqosh (IC1): *bšala d-našə* ‘the cooked food of people’ (Khan 2002a: 207)

Even in such dialects, however, there is a tendency to suffix the particle to the head noun, e. g.

C. Qaraqosh (IC1): *hukizd sawi* ‘the story of my grandfather’

The suffixing of the particle to the head noun is the norm in most dialects. In such dialects the *d* is generally devoiced to *t*:

C. Barwar (IC2): *bronət malka* ‘the son of the king’ (Khan 2008: 397)

J. Arbel (1J2): *babət Yosəf* ‘the father of Joseph’ (Khan 1999: 168)

In some dialects the ending *-at* of the head noun is contracted or elided, especially in phrases expressing kinship relations or inalienable possession, e. g.

C. Barwar (IC2): *brōn mami* ‘the son of my uncle’ (Khan 2008: 400)

In the Jewish trans-Zab dialects on the east periphery of the region the subordinating particle is generally omitted and the head noun has its normal nominal inflectional vowel:

J. Sulemaniyya (IJ2): *šəmma bronā* ‘the name of the boy’ (Khan 2004: 192)

The genitive particle *d* remains prefixed to demonstrative pronouns in the dependent phrase in dialects that normally mark annexation by a *-at* suffix on the head phrase and even in IJ2 dialects like J. Sulemaniyya that do not normally use a genitive particle in annexation. There are grounds for interpreting these as reflecting the bonding of the particle with the pronoun to form a genitive allomorph of the pronoun. This is shown by the fact that it may be combined with a head noun with *-at* in dialects that use this suffix:

- (15) C. Barwar (IC2) (Khan 2008: 399)

'ahwalt-at d-o naša
condition-GEN GEN-that man
'the condition of that man'

- (16) J. Sulemaniyya (IJ2) (Khan 2004: 192)

bela d-oni
house GEN-those
'the house of those (people)'

In some dialects an analytical type of genitive construction is used in which the dependent noun is introduced by an independent genitive particle. This is the case, for example, in J. Arbel (IJ2), which uses the particle *'ot* in this function (< *'o* 3s pronoun + annexation suffix *t*). The head noun in such constructions may optionally have the annexation suffix *-at*:

- (17) a. J. Arbel (IJ2) (Khan 1999: 224)

bšəlmane 'ot- 'Arbel
Muslims GEN-Arbel
'the Muslims of Arbel'

b. *jiran-at 'ot hula 'e*
neighbours-GEN GEN-Jews
'the neighbours of the Jews'

In the Jewish trans-Zab dialects an Iranian *ezafe* particle is occasionally used in annexation constructions, e. g.

J. Sulemaniyya (IJ2): *maktab i hulaye* ‘the school of the Jews’ (Khan 2004: 192)

7. Verbal morphology

The NENA dialects exhibit a root and template type of verb morphology, which is characteristic of Semitic languages. A root is a lexical morpheme consisting of a non-concatenative sequence of radicals. This sequence of radicals is merged with non-concatenative templatic patterns consisting of vowels and, in some cases, additional consonants, that express inflection and derivation of the lexical root. Verbal roots consist of either three (trilateral) or four radicals (quadrilateral).

Inflections of verbs in NENA are formed from a series of inflectional bases. Representative examples of these from the C. Qaraqosh dialect are as follows:

Table 3: Inflectional bases of the verb ‘to open’ in C. Qaraqosh

Stem I

p-θ-x ‘to open’ (trilateral)

Present base:	<i>paθəx</i>
Past base:	<i>pθix</i>
Resultative participle:	<i>pθixa</i>
Imperative:	<i>pθox</i>
Infinitive:	<i>pθaxa</i>

The present base is derived historically from the active participle of earlier Aramaic, and the past base and resultative participle from the erstwhile passive participle. The finite present and past forms of earlier Aramaic (*yiqtol* and *qtal* respectively) have been lost.

NENA preserves some of the derivational verbal morphology of earlier Aramaic, whereby derived forms of verbs are created by morphologically more complex sets of inflectional base patterns. The dialects have either one or two derived sets of inflectional bases. C. Qaraqosh has two derived patterns, which may be designated stem II and stem III, the simple pattern being stem I:

Table 4: Verbal stem II in C. Qaraqosh

q-d-m ‘to present, propose’

Present base:	<i>mqadəm</i>
Past base:	<i>mqudəm</i>
Resultative participle:	<i>mqudma</i>
Imperative:	<i>qadəm</i>
Infinitive:	<i>qadomə</i>

Table 5: Verbal stem III in C. Qaraqosh

<i>r-k-x</i> ‘to soften’	
Present base:	<i>markəx</i>
Past base:	<i>murkəx</i>
Resultative participle:	<i>murkxa</i>
Imperative base:	<i>'arkəx</i>
Infinitive:	<i>'arkoxə</i>

The main function of the derived verbal stems is to express increases of transitivity. In earlier Aramaic there were derived stems that expressed a decrease in the transitivity of the basic stem I, but these have been lost in NENA (Göransson 2015; Khan 2015). C. Qaraqosh preserves the most archaic forms of the inflectional and derivational patterns of the verbs, with the greatest morphological heterogeneity. In most other NENA dialects various degrees of analogical levelling have taken place within the set of inflectional bases of the derivational stems and across derivational stems. This levelling is most advanced in dialects in the northeast of Iraq, in the groups IC3 and IJ2. In many Christian dialects of IC3 and all Jewish dialects of IJ2 the stem II derivation has been lost (Mutzaifi 2004c). In some of the Jewish IJ2 dialects on the northeastern periphery of Iraq the levelling of the vowel patterns across the two surviving derivational stems is total, e. g.

Table 6: Stem I, J. Sulemaniyya

<i>p-l-x</i> I ‘to open’		
	Transitive	Intransitive
Present base:	<i>păləx</i>	<i>păləx</i>
Past base:	<i>pləx</i>	<i>plix</i>
Resultative participle:	<i>pălxə</i>	<i>plixa</i>
Imperative:	<i>păləx</i>	<i>păləx</i>
Infinitive:	<i>păloxe</i>	<i>păloxe</i>

Table 7: Stem II, J. Sulemaniyya

<i>m-rd-x</i> II ‘to boil’		
	Transitive	Intransitive
Present base:	<i>mardəx</i>	<i>mardəx</i>
Past base:	<i>mərdəx</i>	<i>mərdix</i>
Resultative participle:	<i>mərdxa</i>	<i>mərdixa</i>
Imperative:	<i>mardəx</i>	<i>mardəx</i>
Infinitive:	<i>mardoxe</i>	<i>mardoxe</i>

It can be seen that stem I and stem II have identical vowel patterns. In the present, imperative and infinitive the vocalism of stem II has been generalized to stem I. In the past and resultative participle the vocalism of stem II, which is typically causative transitive in function, has been transferred to transitive verbs of stem I. Vice-versa, the original vocalism of the past and resultative participle of stem I is preserved in intransitive verbs and this has been extended also to stem II to express a detransitivized, anticausative.

The present and past inflectional bases are inflected for person and number by two sets of suffixes, referred to here as D-suffixes and L-suffixes, which indicate the grammatical relations of verbal arguments in the clause.

D-suffixes are historically ‘direct’ clitic pronouns agreeing in number, gender and person with the nominative subject of a clause in the original nominative—accusative alignment system of Aramaic. L-suffixes are historically prepositional phrases consisting of the dative preposition *l-* and a pronominal suffix.

The forms of the suffixes in J. Sulemaniyya are as follows:

Table 8: D- and L-suffixes, J. Sulemaniyya

	D-suffixes	L-suffixes
3MS	-Ø	-le
3FS	-a	-la
3PL	-i	-lu
2MS	-et	-lox
2FS	-at	-lax
2PL	-etun	-läxun
1MS	-na	-li
1FS	-an	-li
1PL	-ex	-lan

The D-suffixes are morphologically less marked than the L-suffixes. Historically the third person D-suffixes contain an expression only of number and gender, the person marker of the third person being zero. In the case of the 3ms form, also the number and gender marking is zero. The first and second person D-suffixes have developed from a coalescence of number, gender and person markers. The L-suffixes, on the other hand, all contain pronominal elements that syncretize person, number and gender, as well as a relational element *l-*, which is historically a dative preposition.

The /l/ of the L-suffixes assimilates to a stem-final sonorant /n/ or /r/ in most NENA dialects, e. g. C. Barwar (IC2) *kpinne* ‘he became hungry’ (<*kpin-le*). In some Christian dialects in the northeast of the region this assimilation has been extended to other consonants, e. g. C. Diyana-Zariway (IC3) *ptəxxe* ‘he opened’ (<*ptəx-le*).

In J. Sulemaniyya, in the northeastern periphery of Iraq, there is ergative alignment in clauses with verbs derived from the past base. Clauses with verbs derived from the present base, on the other hand, have nominative–accusative alignment. The following examples show how an intransitive subject (S) aligns with a transitive subject (A) in clauses with present base verbs but aligns with the transitive object (O) in clauses with past base verbs. In each case the common marker is the 3fs D-suffix on the verb. The case-marking of the arguments of the clause is expressed by the cross-referencing D-suffixes and L-suffixes on the verb. Verbs derived from the past base express, in principle, the past perfective:

- (18) J. Sulemaniyya
- Present base
- a. *brat-i barux-e garš-á-lu*
daughter-my friend-PL pull.PRS-D.3FS-L.3PL
'My daughter (A) pulls the friends (O).'
 - b. *brat-i samx-a*
daughter-my stand.PRS-D.3FS
'My daughter (S) stands.'
- Past base
- c. *barux-e brat-i gərš-á-lu*
friend-PL daughter-my pull.PST-D.3FS-L.3PL
'The friends (A) pulled my daughter (O).'
 - d. *brat-i smix-a*
daughter-my stand.PST-D.3FS
'My daughter (S) stood up.'

This type of ergative alignment is found also in NENA dialects over the border in western Iran (Khan, this volume, chapter 4.4, §7.1). As can be seen, there is an inversion of agreement of the two sets of suffixes with grammatical arguments in present base and past base verbs. For present base verbs, a D-suffix cross-references the A argument and an L-suffix cross-references the O argument, whereas in past base verbs, this is reversed, a D-suffix cross-references the O argument and an L-suffix cross-references the A argument. In intransitive clauses with both present base and past base verbs a D-suffix cross-references the S argument:

- (19) Alignment of verbal agreement, according to stem and transitivity
- a. V.PRS—D-suffix(A)—L-suffix(O)
 - b. V.PST—D-suffix(O)—L-suffix(A)
 - c. V.PRS—D-suffix(S)
 - d. V.PST—D-suffix(S)

Constructions such as (18c), in which the A argument is expressed by the historically dative L-suffix, are active ergative constructions rather than passive constructions.⁸ The suffixes that cross-reference the O argument in both present-base constructions (like 18a) and past-base constructions (like 18c) are only present if the O is definite.

The type of alignment of J. Sulemaniyya is, in fact, what Dixon (1994: 71) refers to as split-S rather than canonical ergativity. D-suffixes on the past base occur only when the subject of the intransitive verb is an undergoer of the action. An ‘undergoer’ subject is one that is affected by a change in configurational state or location. When the subject is agentive or is the source of the action (e.g. the source of the production of sound), an intransitive verb has an L-suffix like transitive verbs, e.g. *srax-le* ‘He shouted’ (Khan 2004: 300).

In most NENA dialects in Iraq, however, a hybrid form of alignment is found in verbs formed from past bases. In this type of alignment, which may be called ‘extended ergative’ (Doron and Khan 2012), the marked L-suffixes are used to express both transitive and intransitive subjects, e.g.

- (20) C. Barwar

- a. *xawr-ǎwaθ-i griš-a-la brati-i*
friend-PL-my pull.PST-D.3FS-L.3PL daughter-my
‘My friends pulled my daughter.’
- b. *brati qim-la*
daughter-my rise.PST-L.3FS
‘My daughter rose.’

This type of case-marking should not be analysed as nominative–accusative, since the transitive clauses exhibit some characteristic features of ergativity (Doron and Khan 2012: 230–233). This is likely to have been the more archaic type of construction in NENA. The more canonical type of ergative alignment found in the eastern periphery of NENA in J. Sulemaniyya and Iran appears to be an innovation in NENA, induced by a greater degree of convergence with the ergative constructions of Kurdish (Khan 2016).⁹

The present base, which is historically an active participle, is used in NENA dialects with a variety of functions, in many cases with the addition of auxiliary particles. The unmarked form without auxiliary particles is widely used to express unrealis. Present indicative and future are formed from the present base by means of added particles, e.g.

⁸ For the arguments in favour of this analysis see Doron and Khan (2012) and Khan (2016).

⁹ For more details about ergativity in NENA see Coghill (2016).

(21) C. Barwar

3MS	<i>paθəx</i>	‘(that) he open’ (irrealis)
	<i>paθəx-wa</i>	‘(that) he opened’ (irrealis past)
	<i>'i-paθəx</i>	‘he opens’ (indicative habitual)
	<i>'i-paθəx-wa</i>	‘he used to open’ (indicative habitual past)
	<i>bəd-paθəx</i>	‘he will open’ (future)
	<i>bəd-paθəx-wa</i>	‘he would open’ (future in the past)

The indicative particles that are used in the dialects of the region include *'i-*, as in C. Barwar, and *k-*, e. g. C. Qaraqosh (IC1) *k-paθəx*. The *'i-* is likely to be a reduced form of the copula.

Several dialects in Iraq use a particle with the form *qam-* prefixed to the present base to express the perfective past.¹⁰ In some dialects it has undergone various degrees of phonetic lenition, e. g. C. Barwar *qəm-*, C. Qaraqosh *kəm-*, C. Sulemaniyya *tam-*. This construction was used in place of constructions formed from the past base to avoid using D-suffixes to express pronominal objects on past perfective verbs formed from the past base. This is the case in particular when the pronominal object is first or second person.¹¹ The present base allows the use of L-suffixes to express the direct pronominal object, e. g.

(22) C. Barwar1

- a. *griš-la*
pull.PST-L.3FS
‘she pulled’
- b. *qəm-garš-a-li*
PST-pull.PRS-D.3FS-L.1S
‘she pulled me’

Most NENA dialects do not allow the accumulation of two pronominal object suffixes on ditransitive verbs. If a verb has a pronominal direct object suffix, a pronominal indirect object must, in many dialects, be expressed by an independent prepositional phrase:

(23) C. Barwar

- | | |
|---------------------|---------------|
| <i>yawəl-le</i> | <i>tlal-i</i> |
| give.PRS.3MS-L.3MS | to-1s |
| ‘he gives it to me’ | |

¹⁰ For the historical background of this particle see Fassberg (2015).

¹¹ For the motivation for the repair mechanism see Doron and Khan (2012) and Khan (2016).

In some dialects in the group IC1 on the Mosul plain, such as C. Qaraqosh (Khan 2002a: 143–144) and C. Telkepe (Coghill 2010), and also dialects further east in the IC3 group, such as C. Shaqlawa, present base verbs may express the indirect pronominal object by an L-suffix and the direct pronominal object by a form that resembles an enclitic copula (see below) and can be interpreted as a enclitic pronoun (Khan 2012a). This expression of the pronominal direct object is restricted to the third person:

- (24) C. Qaraqosh
kewi-ləh=ina
 give.PRS.3PL-L.3MS=PRO.3PL
 'They give them to him'

The NENA dialects in Iraq have various types of copula. These include deictic copulas, which draw attention to a referent or to a situation, and neutral copulas which express simple predication. Paradigms of the deictic and the simple copula in C. Qaraqosh and C. Barwar are as follows:

Table 9: Copulas in NENA dialects of Iraq

Deictic copula		Simple copula	
C. Qaraqosh	C. Barwar	C. Qaraqosh	C. Barwar
3MS <i>kilə</i>	<i>hole</i>	= <i>ilə</i>	= <i>ile</i>
3FS <i>kila</i>	<i>hola</i>	= <i>ila</i>	= <i>ila</i>
3PL <i>kina</i>	<i>hole</i>	= <i>ina</i>	= <i>ile</i>
2MS <i>kiyət</i>	<i>hot</i>	= <i>iyət</i>	= <i>iwət</i>
2FS <i>kiyat</i>	<i>hot</i>	= <i>iyat</i>	= <i>iwat</i>
2PL <i>kiyetu</i>	<i>hotu</i>	= <i>iyetu</i>	= <i>iwitu</i>
1MS <i>kiyən</i>	<i>hon</i>	= <i>iyən</i>	= <i>iwən</i>
1FS <i>kiyan</i>	<i>hon</i>	= <i>iyan</i>	= <i>iwən</i>
1PL <i>kiyax</i>	<i>hox</i>	= <i>iyax</i>	= <i>iwəx</i>

The simple copula is generally attached to the predicate as an enclitic, though in some dialects it has a degree of freedom of movement.

A characteristic feature of many of the NENA dialects of Iraq is the expression of the resultative perfect by combining the resultative participle with a copula, e. g.

- C. Qaraqosh *kilə griša* 'he has pulled'
 C. Barwar *grišele* (= *griša + ile*)

Unlike clauses that have past base verb forms, resultative perfect constructions have nominative–accusative alignment in most dialects, with the copula and participle agreeing with the subject in transitive clauses. In (25) J. Sulemaniyya (IJ2)

on the eastern periphery of the region, however, a hybrid type of alignment is found, whereby the copula agrees with the subject but the participle agrees with the object, e. g.

- (25) J. Sulemaniyya (IJ2)

'o-gora baxt-āke gr̥ašta-aw-ye
that-man woman-the pull.PTC.FS-PRO.3FS-COP.3MS
'That man has pulled the woman'

In several dialects in the northeastern sector of the region, resultative perfects are formed by a fossilized, uninflected form of the copula, e. g. *lā* (< 3fs copula *-ila*). This is found, for example, in C. Bədylal (Table 10), as shown in the paradigm below. In the third person, *lā* is used with the participle alone, the distinction between 3ms, 3fs and 3pl being expressed by the gender and number inflection of the participle. In the first and second persons a person marker is required, so an inflected form of the enclitic copula is added to the participle:

Table 10: Resultative perfect in C. Bədylal (IC3)

3MS	<i>lā griša</i>	'he has pulled'
3FS	<i>lā gr̥ašta</i>	'she has pulled'
3PL	<i>lā griše</i>	'they have pulled', etc.
2MS	<i>lā grišewət</i>	
2FS	<i>lā grišewat</i>	
2PL	<i>lā grišütün</i>	
1MS	<i>lā grišewən</i>	
1FS	<i>lā gr̥išewan</i>	
1PL	<i>lā gr̥išewax</i>	

In some Jewish dialects of IJ2, a resultative perfect is formed by combining a fossilized copula with a past base verb form, e. g.

- (26) J. Arbel (IJ2)

lā gr̥eš-le 'he has pulled'
lā gr̥eš-la 'she has pulled'
lā gr̥eš-lu 'they have pulled', etc.

In the J. Rustaqā dialect (IJ2), originally spoken on the northeastern periphery of the region, the resultative perfect exhibits an additional distinctive feature, viz. the inflection of the past base with D-suffixes rather than L-suffixes in intransitive clauses. Intransitive past perfective verbs are inflected with L-suffixes:

Table 11: Resultative perfect in J. Rustaqa (IJ2)

Intransitive past perfective:	<i>qim-la</i> rise.PST-L.3FS 'She rose'
Transitive perfect:	<i>grəš-la</i> pull.PST-L.3FS 'She pulled'
Intransitive perfect:	<i>lā qim-a</i> rise.PST-D.3FS 'She has risen'
Transitive perfect:	<i>lā grəš-la</i> COP pull.PST-L.3FS 'She has pulled'

In many NENA dialects of the region the perfect verbal form is used not only to express the resultative perfect but is also used to express the evidential or an event in the remote past. In some dialects it is used as a narrative form, especially in folktales (Khan 2012b).¹²

The majority of NENA dialects in Iraq express the present progressive by constructions containing the copula. Many combine the copula with the infinitive to perform this function. The original form of this construction contained the locative preposition *b-* 'in', which is still used in some dialects in some circumstances. This seems to have arisen under the influence of Eastern Armenian, which exhibits a similar construction for the progressive (see Martirosyan, this volume, chapter 2.2, §4.2), e. g.

- (27) Eastern Armenian, Ardvin
mnalis im
 stay.INF.LOC COP.1S
 'I am staying'

The preposition *b-* is, however, frequently dropped and this is the regular situation in many dialects, e. g.

- (28) J. Betanure (IJ1)
 'ile bə-graša ~ 'ile graša 'he is pulling'
 'ila bə-graša ~ 'ila graša 'she is pulling'
 'ilu bə-graša ~ 'ilu graša 'they are pulling'

¹² The use of one and the same form for resultatives, evidentials, and as a special narrative form for use in folktales, is clearly reminiscent of the range of functions associated with the verbal suffix *-mIṣ* in Turkish.

This construction is not used in some dialects along the southern periphery of the region. Instead progressive constructions are used in which the copula is combined with a verb form derived from the present base, e. g.

- (29) C. Qaraqosh (IC1)

<i>kilə k-garəš</i>	'he is pulling'
<i>kila k-garša</i>	'she is pulling'
<i>kina k-garši</i>	'they are pulling'

In the eastern sector of the region many dialects use a fossilized uninflected form of the copula in progressive constructions. In some cases this is combined with a present base verb form, e. g. J. Arbel *lā gärəš* 'he is pulling', J. Dobe *nā gärəš*, C. Xərpə *ho k-garəš*, Shosh u-Sharmən *nə-k-garəš*, C. Bədylal *ma-k-garəš* (Mutzafi 2004c; Coghill 2008). In some dialects the fossilized copula is combined with the infinitive. Since the infinitive is not inflectable for gender, number or person, these must be marked by inflected enclitic copulas in all slots of the paradigm beyond the 3ms, e.g.

Table 12: Progressive constructions with copula, J. Rustaqa (IJ2) (Khan 2002b: 407)

3MS	<i>lā garoša</i>	'he is pulling'
3FS	<i>lā garoše=la</i>	'she is pulling'
3PL	<i>lā garoše=lu</i>	'they are pulling', etc.
2MS	<i>lā garoša=wet</i>	
2FS	<i>lā garoša=wat</i>	
2PL	<i>lā garoša=wetu</i>	
1MS	<i>lā garoša=wena</i>	
1FS	<i>lā garoša=wan</i>	
1PL	<i>lā garoša=wex</i>	

In C. Ankawa (IC3) the present progressive is expressed by combining the particle *də-* with present base verbs, e. g. *də-k-garəš* 'he is pulling'. This particle may be a phonetic variant of the fossilized copula. Alternatively it may be identified as a subordinating particle which presents the verb as complement of a zero copula '(it is) that he pulls' (Borghero 2015). Indeed, the constructions with an explicit invariable form of the copula described above may have a similar cleft syntactic structure, e. g. J. Arbel *lā gärəš* 'it is (that) he pulls'.

In some dialects the imperative is strengthened by prefixing the particle *da-*, *də-*, which is a loan from Kurdish, e. g. J. Arbel *da-holli!* 'give me!' (Khan 1999: 282–283), C. Barwar *də-plut!* 'go out (sing.)!' (Khan 2008: 179). C. Barwar also uses suffixes of Turkic origin added to the imperative, e. g. *plut-gən!* 'go out (sing.)!' (Khan 2008: 179).

8. Syntax of basic clauses

8.1. Copula clauses

The default position for the simple basic copula in predication is after the predicate, generally as an enclitic. This applies to both equative clauses, identifying the referent of the subject with that of a nominal in the predicate, or ascriptive clauses, which ascribe a property to the referent of the subject, e. g.

(30) C. Barwar (IC2)

- a. *mamux šwawan=ile*
uncle.your neighbour.our=COP.3MS
'Your uncle is our neighbour.'
- b. *mamux kpinele (< kpina=ile)*
uncle.your hungry=COP.3MS
'Your uncle is hungry.'

In many dialects the copula is moveable. In C. Barwar, for example, it is sometimes placed before the predicate in equative clauses, e. g.

(31) C. Barwar

- mamux 'ile šwawan*
uncle.your COP.3MS neighbour.our
'Your uncle is my neighbour.'

In such dialects the copula may be cliticized to components of the clause that are in narrow focus, e. g. in (32), in which the referent of the subject of the clause is in narrow focus, the remainder of the clause being presuppositional. The narrow focus, represented by upper case below, is expressed prosodically by the nuclear stress of the intonation group.¹³

(32) C. Barwar

- a. *màmux=ile šwáwan.*
uncle.your=COP.3MS neighbour.our
'YOUR UNCLE is our neighbour.'
- b. *màmux=ile kpína.*
uncle.your=COP.3MS hungry
'YOUR UNCLE is hungry.'

In some dialects the positioning of the copula is more rigidly fixed and regularly occurs as an enclitic on the predicate irrespective of the type of the predication

¹³ In the examples the intonation group boundary is represented by |, the nuclear stress of the intonation group by a grave accent (̀) and non-nuclear stresses by acute accents (́).

or the information structure of the clause. This applies in particular to the Jewish dialects in IJ2, e. g.

(33) J. Sulemaniyya

- a. *šwáwan tujár=ye.*¹
neighbour.our merchant=COP.3MS
'Our neighbour is a merchant.'
- b. *šwáwan jəhyà=ye.*¹
neighbour.our tired=COP.3MS
'Our neighbour is tired.'
- c. *šwàwan tujár=ye.*¹
'OUR NEIGHBOUR is a merchant.'
- d. *šwàwan jəhyá=ye.*¹
'OUR NEIGHBOUR is tired.'

The copula is negated by prefixing to it the negative particle *la*, e. g. C. Barwar *lele* 'he is not' (<*la* + *ile*), J. Sulemaniyya *la-y* 'he is not' (<*la-y*). In some cases, the negative copula has a different stem from the positive copula, e. g. J. Arbel (IJ2) *lewe* 'he is not' (cf. *ile* 'he is'), C. Bədylal (IC3) *liwe* 'he is not' (cf. *ile* 'he is').

In most dialects of the region the negated copula is placed before the predicate:

(34) C. Barwar

- bábi lélé làxxa.*¹
father.my NEG.COP.3MS here
'My father is not here.'

In the Jewish dialects of the IJ2 group, however, it is placed after the predicate, e. g.

(35) J. Sulemaniyya

- tatí laxxá là-y.*¹
father.my here NEG-COP.3MS
'My father is not here.'

8.2. Verbal clauses

The predominant order of the core arguments of intransitive verbal clauses (S and V) is SV in all dialects. The core arguments of transitive clauses (A, V and O) is AVO in most dialects, but AOV in the Jewish trans-Zab dialect group (1J2), e. g.

(36) C. Barwar (IC2)

- a. *šwawan bəd- 'axəl-Ø laxma*
neighbour.our FUT-eat.PRS-D.3MS bread
'Our neighbour will eat bread.'

- J. Sulemaniyya (IJ2)
- b. *šwawan* *ləxma* *k-xəl-Ø*
 neighbour.our bread IND-eat.PRS-D.3MS
 ‘Our neighbour will eat bread.’

A prepositional phrase expressing an indirect object or some other complement of the verb is, however, normally placed after the verb in all dialects:

- (37) C. Barwar
- a. *bəd-taʃq-əx* *b-šwawan*
 FUT-meet.PRS-D.1PL on-neighbour.our
 ‘We shall meet our neighbour.’
- J. Sulemaniyya
- b. *doq-ex* *b-šwawan*
 hold.PRS-D.1PL on-neighbour.our
 ‘We shall hold onto our neighbour.’

A nominal without a preposition that expresses the goal of a verb of movement is normally placed after the verb in all dialects, including those of the IJ2 group, e. g.

- (38) C. Barwar
- a. *zil-lan* *bəθa*
 go.PST-L.1PL home
 ‘We went home.’
- J. Sulemaniyya
- b. *zil-ex* *bela*
 go.PST-D.1PL home
 ‘We went home.’

In all dialects the predominant order of clause components may be changed for pragmatic purposes.

Verbal clauses are negated by the negative particle *la*, e. g.

- (39) C. Barwar (IC2)
- šwawan* *la xil-le* *laxma*
 neighbour.our NEG-eat.PST-L.3MS bread
 ‘Our neighbour did not eat bread.’

The future particle is not used when the verb is negated, e. g. C. Barwar *bəd-’axəl* ‘he will eat’, *la-’axəl ~ la-y-axəl* ‘he will not eat’. The verb in the second of the negated alternatives (*y-axəl*) has the habitual prefix.

In all dialects except some in the southwestern periphery of the region on the Mosul plain (the IC1 group) the imperative form may be negated, e. g. C. Barwar (IC2) *la plu!* ‘Do not go out.’ In C. Qaraqosh (IC1) a prohibition cannot be expressed by negating an imperative, but rather the negative particle is com-

bined with an unrealis verb form, e. g. *la paltat!* ‘Do not go out!’ (Khan 2002a: 351).

In dialects that express the progressive or perfect by a prefixed fossilized copula particle *lā*, the negator is placed after this particle, e. g.

- (40) J. Arbel (Khan 1999: 115)

lā la-hqe-li
COP NEG-speak.PST-L.1s
'I have not spoken.'

In the Jewish on the eastern periphery dialects when *la* negates a perfect formed by the combination of a resultative participle and a copula auxiliary, the negator is placed before the participle rather than the auxiliary, e. g.

- (41) J. Sulemaniyya (Khan 2004: 105)

- a. *la smiqa-y*
NEG reddens.PTC-COP.3MS
'It has not become red.'
- b. *smoqa la-y*
red NEG-COP.3MS
'It is not red.'

9. Syntax of clause linkage

9.1. Coordinating linkage

The coordinative linkage of main clauses may be asyndetic or marked by a particle. When marked by a particle, this is generally either (i) a clausal connective particle or (ii) a particle marking topical prominence.

The first strategy is performed by the Aramaic connective particle *w* or the loaned Arabic particle *fa-*. The usual practice is to place the particle before the last in a chain of sequential clauses, e. g.

- (42) C. Qaraqosh (IC1)

r̥aq-lə baθr-èhə, kəm-madər-Ø-hə
run.PST-L.3MS after-them, PST-return.PRS-D.3MS-L.3PL
u-θi-lə l-bəθə
and-come.PST-L.3MS to-house
'He ran after them, returned them and came home.' (Khan 2002a: 463)

In general, the particle *fa* marks a greater boundary in the discourse than does *w*. In a narrative, speakers use *fa* as a means of signalling the onset of the main stages in the progression of the events (Khan 2002a: 465, 2008: 918).

Particles marking topical prominence that are used in clausal connection include the Aramaic particle *'ap* and the Kurdish enclitic *-ži* or phonetic variants

thereof (e.g. *-iš*, all glossed below as TOP). This is typically combined with a clause initial subject, which has an identifiable referent, e.g.

- (43) C. Barwar (IC2), (Khan 2008: 920)

<i>zilte-la</i>	<i>tóṭo </i>	<i>már-a</i>	<i>màlka ...</i>
go.PTCP.-COP.3FS	old.woman	said.PST-L.3FS	king
<i>'áp- 'awwa</i>	<i>šaqál-Ø-la</i>	<i>ðá-kista</i>	<i>qa-dáy. </i>
TOP-he	take.PRS-D.3MS-L.3FS	one-purse	for-OBL.3FS

‘The old woman went and said “Oh King ...” Then, he takes a purse for her.’

- (44) J. Sulemaniyya (IJ2) (Khan 2004: 401)

<i>xól-la</i>	<i>hyè-Ø-wa. </i>	<i>bába</i>	<i>bràt-áke-š </i>
indeed-not	come.PST-D.3MS-back	father	daughter-the-TOP
<i>m-şálma komùl-ew </i>	<i>mad-hítwa-le</i>	<i>zbàn-ne. </i>	
from-face blackness-his	what-EXIST.PST-L.3MS	sell.PST-L.3MS	

‘He indeed he did not come back. The father of the girl, out of his shame sold what he had.’

The use of topical prominence particles to mark clausal connection is not used in C. Qaraqosh and many dialects of IC1 on the western periphery of the region.

9.2. Subordinating linkage

The general subordinating particle *d* inherited from earlier Aramaic is widely used to introduce subordinate clauses of various types in the Christian dialects of Iraq and the IJ1 group of Jewish dialects. These include the following:

9.2.1. Relative clauses

Some dialects preserve the archaic construction of earlier Aramaic in which the particle is attached to the beginning of the modifying relative clause:

- (45) C. Qaraqosh (IC1), (Khan 2002a: 475)

<i>'áwa</i>	<i>qéṭma</i>	<i>d-k-pàyəš-Ø-wa.</i>
that	ash	REL-IND-remain.PRS-D.3MS-PST

‘the ash that was left’

In many dialects the particle is suffixed to the head noun and devoiced:

- (46) C. Barwar (IC2), (Khan 2008: 952)

<i>'e-brat-ət</i>	<i>maluš-i-wa-la</i>	<i>jullə</i>
that-girl-REL	dress.PRS-D.3PL-PST-L.3FS	clothes

‘the girl whom they dressed in clothes’

In the Jewish IJ2 dialect J. Arbel the particle generally combines with an anaphoric pronoun rather than being attached directly to the head noun:

- (47) J. Arbel (IJ2), (Khan 1999: 388)

kré 'ot-loš-i-lu
 suits DEM.REL-wear.PRS-D.3PL-L.3PL
 'the suits that they wore'

9.2.2. Clausal complements of adverbs and prepositions

- (48) C. Qaraqosh (IC1), (Khan 2002a: 488)

'émma d-g-ya 'l-i-wa l-íta
when COMP. IND. enter.PRS. D. 3.PL. PST. to church

when COMP-IND-enter.PRS-B.E
šaly i wa šan̥e |

sulx-t-wa *carox*
take off PBS D 3_{PL} PST shoes

‘When they entered a church, they took off their shoes.’

In many dialects the particle is suffixed to the adverbial head:

- (49) C. Barwar (JC2). (Khan 2008: 982)

C. Barwar (TCZ), (Khan 2008: 982),
 'íman-t v-əθ-é-wa dwára aúrbo

when-COMP IND-come PBS-D 3PL-PST pasture near

l-màθq¹ *v-asq-i-wq* *berive*

to-village IND-ascend.PRS-D.3PL-PST milkmaids

'When they came to the pasture land near the village, milkmaids would come up.'

9.2.3. Irrealis complements of verbs

- (50) C. Oaraqosh (IC1). (Khan 2002a: 505)

'áxni g-bá'-ax d-šák-ax l-àna.

we IND-want.PRS-D.1PL COMP-complain.PRS-D.1PL about-them

‘We want to complain concerning them.’

9.2.4. Factual complements of verbs

- (51) C. Qaraqosh (IC1), (Khan 2002a: 506)

ðá'-la *d-íla* *niyát-tə* *xrùta.*

know.PST-L.3FS COMP-COP.3FS intention-their bad

‘She knew that their intention was bad.’

In some dialects the *d* is occasionally replaced by a particle from a contact language, as in the following factual complement clause:

- (52) C. Qaraqosh (IC1), (Khan 2002a: 506)

<i>k-áð-i</i>	<i>bə'ánnahu</i>	<i>k-ína</i>	<i>tíwə</i>	<i>'áñə</i>
IND-KNOW.PRS-D.3PL	COMP	IND-COP.3PL	settled	DEM

suráya.!

Christians

‘They know that (= Arabic *bi'annahu*) the Christians have settled there.’

The replacement of *d* by loaned particles is the norm in the IJ2 dialects on the eastern periphery of the region:

- (53) J. Sulemaniyya (IJ2), (Khan 2004: 414–415)

<i>bába</i>	<i>'ó-brona</i>	<i>ga-ləbl-á-le</i>	<i>hají-ye-le.</i> !
father	DEM-boy	REL-take.PST-D.3FS-L.3MS	haji-COP.PST-L.3MS

‘The father of the boy who (= Kurdish *ga/ka*) took her away was a *hají*.’
(relative clause)

- (54) J. Sulemaniyya (IJ2), (Khan 2004: 440)

<i>kăy-én-wa</i>	<i>ga-'</i> ó	<i>bratá</i>	<i>'əl-do-bróna</i>
know.PRS-D.3PL-PST	COMP-DEM	girl	OBJ-OBL.DEM-boy

gb-à.!

love.PRS-D.3FS

‘They knew that the girl loved the boy.’ (complement clause)

In some dialects a subordinating particle is omitted in various constructions. There is a greater tendency for this in direct complement clauses of verbs (irrealis and realis) and relative clauses. There is also a geographical correlation, in that asynthetic direct complement constructions are more common in the dialects of the central and eastern peripheries than the IC1 group.

The original Aramaic conditional particle *'ən* is used in the protasis of conditional constructions in many dialects of the region, e. g.

- (55) C. Barwar (IC2), (Khan 2008: 1004)

<i>'ən-pálš-əx</i>	<i>mən-dáni</i>	<i>qaṭl-i-lən.</i> !
if-fight.PRS-D.1PL	with-OBL.DEM.PL	kill.PRS-D.3PL.-L.1PL

‘If we fight with them, they will kill us.’

In the IJ2 dialects the particle is regularly replaced by the Kurdish conditional particle *'agar*:

- (56) J. Sulemaniyya (IJ2), (Khan 2004: 436)

<i>'ágar</i>	<i>šabbát</i>	<i>lá</i>	<i>doq-étun,</i> !	<i>xafàd</i>	<i>kəxl-i.</i> !
if	sabbath	NEG	keep.PRS-D.2PL	trouble	eat.PRS-D.3PL

‘If you do not keep the Sabbath, they will be in trouble.’

In some of the dialects on the Mosul plain, the conditional particle is replaced by an Arabic conditional particle:

- (57) C. Qaraqosh (IC1), (Khan 2002a: 496)

<i>kān</i>	<i>k-šaql-i-hə</i>	<i>dàha,¹</i>	<i>'əm-madár-Ø-hə</i>
if	IND-take.PRS-D.3PL-L.3PL	now	FUT-return.PRS-D.3MS-L.3PL
<i>m-bàθər.</i> ¹	from-after		

‘If (= Arabic *kān*) they take them now, he will bring them back.’

10. Concluding remarks

In general, the dialects in the IC1 group exhibit greater archaism than other NENA dialects in Iraq. This applies especially to C. Qaraqosh on the southwestern periphery. This archaism is reflected, for example, in the preservation of interdental consonants and also pharyngeal consonants in some contexts and heterogeneity in the morphology of paradigms of pronouns and verbal forms. Also some aspects of syntax in IC1 can be regarded as representing archaic features of NENA, such as the extended ergative alignment and predominant placement of the direct object argument after the verb. Dialects of other groups exhibit various degrees of innovation in these features, the most innovative being the trans-Zab Jewish dialects in group IJ2. One factor in the archaism of the IC1 group may have been the greater exposure of these dialects to Arabic. In many of the archaic features listed above, for example, parallels can be found in Arabic dialects, which may have had a conservative influence. The innovations of other dialect groups are no doubt due to a greater exposure to Kurdish rather than to Arabic. In the most innovative group, IJ2, this convergence with Kurdish included the development of canonical ergative alignment in dialects in the eastern periphery of Iraq such as J. Sulemaniyya.

The lexicon of the NENA dialects of the region has been extensively influenced by the languages with which they have been in contact. This applies even to the morphologically and syntactically archaic dialects of the IC1 group. Kurdish is a major source of loanwords. This applies in particular to the Christian dialects of the IC2 group and the Jewish dialects. There is an old layer of Kurdish loanwords also in the IC1 group. In the IC1 group numerous Arabic words are used by speakers. Most of these, however, are recent loans, as demonstrated by the fact that most have not been morphologically adapted. Many, indeed, are likely to be the result of code-switching rather than lexical borrowing. The lexicon of all dialects also has some loanwords of Turkic origin, originating mainly either through contact with Turkoman or Ottoman Turkish. In the trans-Zab Jewish dialects calques of Kurdish phrasal verbs containing a noun and a light verb are particularly common, e. g. J. Arbel *bāng 'ol* ‘he invites’ (literally: he makes invitation; cf. Kurdish *bang kirdin*) (Khan 1999: 109).

11. Glossed text sample

C. Barwar dialect (Khan 2008: 1568)

The scorpion and the snake

(^l indicates an intonation group boundary. A grave accent ([˘]) indicates the nuclear stress in the intonation group. Acute accents (^ˊ) indicate non-nuclear stresses in intonation groups).

- (58) *'iθ-wa xa-nàša^l 'áwwa zíl-le zrà'a.^l*
 EXIST-PST one-man he go.PST-L.3MS cultivate.INF
zàrye=le xàčča^l, xàrθa,^l
 cultivate.PTCP.MS=COP.3MS little then
brōn-dìye^l ðé-le šánθ-e dìye,^l
 son-OBL.3MS come.PST-L.3MS sleep-his OBL.3MS
qəm-madmàx-Ø-le.^l
 PST-put.to.sleep.PRS-D.3MS-L.3MS
'iθ-wa xáčča míya slàya,^l šaqiθa.^l
 EXIST-PST little water descend.PROG channel
qəm-maθé-Ø-le támá.^l dmíx-le kəs-dày
 PST-bring.PRS-D.3MS-L.3MS there sleep.PST-L.3MS by-OBL.DEM.FS
šaqiθa.^l 'o-xéna zràye=le.^l
 channel DEM-other cultivate.PROG=COP.3MS
xárθa 'áθye=le xázyele ína
 then come.PTCP.MS=COP.3MS see.PTCP.MS=COP.3MS behold
l-ay-gōt-hòdəx t-šaqiθa^l 'iθ-wa ða-pàqqe.^l
 on-DEM.FS-side-farther GEN-channel EXIST-PST one.F-frog
 ‘There was a man, who went to cultivate. He cultivated a little, then his son became sleepy and he laid him down to sleep. There was some water flowing, a channel. He took him there. He went to sleep by that channel while the other one (his father) was cultivating. Then he came back and he saw that on the farther side of the channel there was a frog.’
- (59) *m-ày-gota^l pàqqe šwír-ra l-a-gōt-hòdəx*
 from-DEM.FS-side frog jump.PST-L.3MFS to-DEM.FS-side-farther
t-šaqiθa.^l klè-la,^l síq-la
 GEN-channel stand.PST-L.3FS ascend.PST-L.3MS
ðá 'aqárwa xás-a dìya.^l
 one.F scorpion back-3FS OBL.3FS
síq-la xás-a dìya^l
 ascend-L.3FS back-3FS OBL.3FS
'u-mášur-á-la b-a-gōt-míya
 and-transfer.PRS-D.3FS-L.3FS at-DEM.3FS-side-water

'áθya lāxxa.¹ brōn-do-náša
 come.IRR.D.3FS here son-OBL.DEM.MS-man
 dmixe=le támma,¹ 'aqérwa slé-la
 sleep.PTCP.MS=COP.3MS there, scorpion descend.PST-L.3FS
 m-xáš-t pàqqe,¹ ðéla wír-ra
 from-back-GEN frog come.PST-L.3FS enter.PST-L.3FS
 xo-réšə do-yala.¹
 under-head OBL.DEM.MS-child

'From one side the frog jumped to the other side of the channel. It stood still and a scorpion mounted onto its back. It mounted its back and it (the frog) took it across to this side of the water so that it came here. The son of that man was sleeping here. The scorpion got down from the back of the frog and went underneath the head of the child.'

- (60) 'u-šqíl-le xá-kepa báb-e diye,¹
 and-take.PST-L.3MS one-stone father-3MS OBL.3MS
 bába do-yala,¹ már-e qatł-án-na
 father OBL.DEM.MS-boy say.PST-L.3MS kill.PRS-D.1MS-L.3FS
 yáha 'aqérwa diya.¹ áyya pàqqe zíl-la
 DEM.FS scorpion now. DEM.FS frog go.PST-L.3FS
 muθy-á-la 'aqérwa m-ay-göt-hódax miya.¹
 bring.PST-D.FS.-L.3FS scorpion from-DEM.3FS-side-farther water
 muθy-á-la láxxa ta-t-nes-á-le
 bring.PST-D.FS.-L.3FS here to-COMP-bite.IRR-D.3FS-L.3MS
 bròn-i,¹ brón-i màyəθ-Ø.¹ qatł-án-na
 son-my son-my die.PRS-D.3MS kill.PRS.-D.1MS-L.3FS
 yáha 'aqérwa.¹ xárθa xá-ga xéta klè-le.¹
 DEM.FS. scorpion then one-time another stand.PST-L.3MS
 'His father, the father of the boy, took a stone and said "I'll kill that scorpion right now. This frog has brought the scorpion from the other side of the water. It has brought it here to bite my son, my son will die. I shall kill that scorpion." Then he refrained.'

- (61) már-e škíra šámm-ux ya- 'álaha.¹
 say.PST-L.3MS praised.PTCP name-your VOC-God
 'án- 'áti báy-æt bròn-i màyəθ-Ø
 if-you want.PRS-D.2MS son-my die.IRR-D.3MS
 b-áyya 'úrxa,¹ 'oð-át-la 'áyya pàqqe bålam,¹
 in-DEM.FS way make.IRR-D.2MS-L.3FS DEM frog boat
 'ázəl-Ø m-day-góta xéta šaqíθa,¹
 go.IRR-D.3MS from-OBL.DEM.FS other.FS channel
 šaql-á-la 'aqérwa l-xáš-a diya,¹
 take.IRR-D.3FS-L.3FS scorpion on-back-3FS OBL.3FS

*máθy-a l-àyya gota, |
 bring.IRR-D.3FS to-DEM.FS side
 'u- 'áθy-a tla-t-nes-á-le brón-i màyəθ-Ø, |
 and-come.IRR-D.3FS to-COMP-bite.IRR-D.3FS-L.3MS son-my die.IRR-D.3MS
 xláp-ux ya- 'álaha. | hár 'əjbóna diyux=ile. |*

sake-your voc-God just will OBL.2MS=COP.3MS

'He said "Praise be your name, oh God, if you will that my son die by this means, that you make this frog into a boat to go from the other side of the channel and take a scorpion on its back, bringing it to this side, for it to bite my son that he should die, let it be as you wish (Literally: It is for your sake) oh God, it is your will."'

- (62) *rapé-Ø-le 'o-kèpa| là qəm-qatál-Ø-la*
throw.PRS-D.3MS-L.3MS DEM.MS-stone NEG PST-kill.PRS-D.3MS-L.3FS
'aqárwa. | 'o-xiš-le| píš-le
scorpion DEM.MS-go.PST-L.3MS remain.PST-L.3MS
zráya zráya. | xá-palgət sà'at, |
cultivate.PROG cultivate.PROG one-half-GEN hour
xá-sa'at xèna, | áw lèle ráhat. |
one-hour other DEM.MS NEG.COP.3MS comfortable
már-e brōn-díyi hár hóle míθa. |
say.PST-L.3MS son-OBL.1s already DEIC.COP.3MS dead.MS
'aqárwa nistə=l-le. | zilə=le|
scorpion bite.PTCP.FS=COP.3FS-L.3MS go.PTCP.MS=COP.3MS
mùklə=l-la táwre t-wéwa zráya. |
stop.PTCP.MS=COP.3MS-L.3PL oxen REL-PST.COP.3PL cultivate.PROG
šqíl-le brón-e diye| már-e
take.PST-L.3MS son-3MS OBL.3MS say.PST-L.3MS
šaql-áñ-na jánáze diye| nabl-áñ-na bëθa. |
take.PRS-D.1MS-L.3FS coffin OBL.3MS take.PRS-D.1MS-L.3FS home
 'He threw away the stone and did not kill the scorpion. He went off and continued to cultivate. Half an hour, an hour later, he was not comfortable. He said "My son is already dead. The scorpion has bitten him". He went and stopped the oxen, which were ploughing. He took his son and said "I'll take his coffin and carry him home".'

- (63) *dárya 'ið-e b-brón-diye. | brón-diye*
put.PTCP.MS hand-3MS on-son-OBL.3MS son-OBL.3MS
qím-wa-le. | már-e ha-báb-i
rise.PST-PST-L.3MS say-L.3MS VOC-father-my
mò báy-ət? | már-e čù-məndi báy-ən. |
what want.PRS.-D.2MS say.PST-L.3MS NEG-thing want.PRS-D.1MS

<i>pr̥iq-li</i>	<i>t-áz-əx</i>	<i>bèθa.</i>
finish.PST-L.1s	FUT-go.PRS-D.1PL	home
<i>xír-re</i>	<i>xo-réšə</i>	<i>do-yàla.</i>
look.PST-L.3MS	under-head	OBL.DEM.MS-boy EXIST-PST
<i>xá-boy-ət</i>	<i>xùwwe.</i>	<i>xá-xuwwe plítə</i>
one-hole-GEN	snake	one-snake leave.PTCP there want-D.3MS
<i>nayás-Ø-le</i>	<i>ṭla-brōn</i>	<i>dó fallāh.</i>
bite-D.3MS-L.3MS	ACC-son	OBL.DEM.M peasant DEM.FS scorpion
<i>θìθa!</i>	<i>gu-réšə</i>	<i>do-bòya!</i>
come.PTCP.FS	in-head	OBL.DEM.MS-hole
<i>q̥iltə=l-le,</i>	<i>qəm-qaq̥l-á-le</i>	<i>'o-xùwwe.</i>
kill.PTCP.FS=COP.3FS-L.3MS	PST-kill.PRS-D.3FS-L.3MS	DEM.MS-snake
<i>már-e</i>	<i>'ən-'ána haw-én-wa</i>	<i>q̥ilə-lla</i>
say.PST-L.3MS	if-I-be.IRR-D.1MS-PST	kill.PTCP-L.3FS
<i>'ăyáha</i>	<i>'aq̥ərwa,</i>	<i>'áwwa xúwwe t-awé-Ø-wa</i>
DEM.FS	scorpion	DEM.3MS snake FUT-be.PRS-D.3MS-PST
<i>q̥ila</i>	<i>ṭla-brōn-i.</i>	<i>šq̥il-le brōn-diye!</i>
kill.PTCP.MS	ACC-son-my	take.PST-L.MS son-OBL.3MS
<i>zíl-le</i>	<i>bèθa.</i>	
go.PST-L.3MS	home	

'He put his hand on his son, his son had woken. He said "Ah, father, what do you want?" He said "I don't want anything. I have finished. Let's go home." He looked under the head of the boy. There was a snake's hole. A snake had come out of there wanting to bite the son of that peasant. The scorpion had come to the opening of the hole and killed it, it killed the snake. He said "If I had killed that scorpion, that snake would have killed my son". He took his son and went home.'

Abbreviations

Abbreviations not found in the Leipzig glossing conventions include:

D = D-suffix (i. e. direct, nominative pronominal suffix)

L = L-suffix (i. e. oblique pronominal suffix)

References

- Aloni, Oz. 2014. *The Neo-Aramaic speaking Jewish community of Zakho: A survey of the oral culture*. Saarbrücken: Lap Lambert.
- Arnold, Werner. 2002. Velarisieretes R und retroflexes R im neuostaramäischen Dialekt von Baratle. *Mediterranean Language Review* 14. 1–8.
- Borghero, Roberta. 2015. The present continuous in the Neo-Aramaic dialect of 'Ankawa and its areal and typological parallels. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 187–206. Piscataway, NJ: Gorgias Press.
- Coghill, Eleanor. 2003. *The Neo-Aramaic Dialect of Alqosh*. Cambridge: University of Cambridge dissertation.
- Coghill, Eleanor. 2008. Some notable features in North-Eastern Neo-Aramaic dialects of Iraq. In Geoffrey Khan (ed.), *Neo-Aramaic Dialect Studies*, 91–104. Piscataway, NJ: Gorgias Press.
- Coghill, Eleanor. 2010. Ditransitive constructions in the Neo-Aramaic dialect of Telkepe. In Andrei L'vovich Mal'chukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook*, 221–42. Berlin & Boston: De Gruyter Mouton.
- Coghill, Eleanor. 2016. *The rise and fall of ergativity in Aramaic: Cycles of alignment change*. Oxford studies in diachronic and historical linguistics 21. Oxford: Oxford University Press.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge studies in linguistics 70. Cambridge: Cambridge University Press.
- Doron, Edit & Geoffrey Khan. 2012. The typology of morphological ergativity in Neo-Aramaic. *Lingua* 122. 225–240.
- Duval, Rubens. 1896. Notice sur les dialectes néo-araméens. *Mémoires de la Société de Linguistique de Paris* 9. 125–135.
- Fassberg, Steven. 2015. The origin of the periphrastic preterite Kəm/qam-Qāṭelle in North-Eastern Neo-Aramaic. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 172–186. Piscataway, NJ: Gorgias Press.
- Fiey, Jean Maurice. 1965. *Assyrie chrétienne: Contribution à l'étude de l'histoire et de la géographie ecclésiastiques et monastiques du nord de l'Iraq*, 2 vols. Recherches sublées sous la direction de l'Institut de Lettres Orientales de Beyrouth, T. 22–23, 42. Beyrouth: Imprimerie catholique.
- Fox, Samuel E. 2008. North-Eastern Neo-Aramaic and the Middle Aramaic Dialects. In Geoffrey Khan (ed.), *Neo-Aramaic Dialect Studies*, 1–18. Piscataway, NJ: Gorgias Press.
- Gaunt, David. 2006. *Massacres, resistance, protectors: Muslim-Christian relations in Eastern Anatolia during World War I*. Piscataway, NJ: Gorgias Press.
- Göransson, Kathrin. 2015. Causative-inchoative alternation in North-Eastern Neo-Aramaic. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*. Piscataway, NJ: Gorgias Press.
- Greenblatt, Jared R. 2011. *The Jewish Neo-Aramaic dialect of Amadiya*. Studies in Semitic Languages and Linguistics, v. 61. Leiden & Boston: Brill.
- Hoberman, Robert. 1988. The history of the Modern Aramaic pronouns and pronominal suffixes. *Journal of the American Oriental Society* 108. 557–575.
- Hoberman, Robert. 1989. Parameters of emphasis: Autosegmental analyses of pharyngealization in four languages. *Journal of Afroasiatic Languages* 2(1). 73–98.

- Jastrow, Otto. 1990. *Der arabische Dialekt der Juden von 'Agra und Arbīl*. Semitica Viva 5. Wiesbaden: Harrassowitz.
- Kahn, Margaret. 1976. *Borrowing and variation in a phonological description of Kurdish*. Michigan: University of Michigan dissertation.
- Khan, Geoffrey. 1999. *A Grammar of Neo-Aramaic: The dialect of the Jews of Arbel*. Leiden & Boston: Brill.
- Khan, Geoffrey. 2002a. *The Neo-Aramaic dialect of Qaraqosh*. Studies in Semitic Languages and Linguistics 36. Leiden & Boston: Brill.
- Khan, Geoffrey. 2002b. The Neo-Aramaic dialect of the Jews of Rustaqa. In Werner Arnold & Hartmut Bobzin (eds.), *Sprich doch mit deinen Knechten Aramäisch, Wir verstehen es! 60 Beiträge zur Semitistik, Festschrift für Otto Jastrow zum 60. Geburtstag*, 395–410. Wiesbaden: Harrassowitz.
- Khan, Geoffrey. 2004. *The Jewish Neo-Aramaic dialect of Sulemaniyya and Halabja*. Studies in Semitic Languages and Linguistics, 44. Leiden & Boston: Brill.
- Khan, Geoffrey. 2007. The North-Eastern Neo-Aramaic dialects. *Journal of Semitic Studies* 52(1): 1–20.
- Khan, Geoffrey. 2008. *The Neo-Aramaic dialect of Barwar*. 3 vols. Leiden & Boston: Brill.
- Khan, Geoffrey. 2009. *The Jewish Neo-Aramaic dialect of Sanandaj*. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2012a. Remarks on the historical development of the copula in Neo-Aramaic. In Federico Corriente, Gregorio del Olmo Lete, Ángeles Vicente & Juan-Pablo Vita (eds.), *Dialectology of the Semitic languages. Proceedings of the IV meeting on comparative Semitics Zaragoza 11/6–9/2010*, 25–31. Sabadell: AUSA.
- Khan, Geoffrey. 2012b. The evidential function of the perfect in North-Eastern Neo-Aramaic dialects. In Rebecca Hasselbach & Na'ama Pat-El (eds.), *Language and nature. Papers presented to John Huehnergard on the occasion of his 60th birthday*, 219–228. Studies in Ancient Oriental Civilization 67. Chicago: University of Chicago.
- Khan, Geoffrey. 2013. Phonological emphasis in North-Eastern Neo-Aramaic. In Jean Léonard & Samia Naïm (eds.), *Base articulatoire arrière. Backing and backness*, 111–132. Munich: Lincon Europa.
- Khan, Geoffrey. 2015. Causative constructions in Neo-Aramaic (Christian Urmi dialect). In Lutz Edzard (ed.), *Arabic and Semitic linguistics contextualized: A Festschrift for Jan Retsö*, 506–530. Wiesbaden: Harrassowitz.
- Khan, Geoffrey. 2016. Ergativity in Neo-Aramaic. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *Oxford Handbook of Ergativity*, 873–899. Oxford: Oxford University Press.
- Khan, Geoffrey. 2017. Studies in the lexicon of Neo-Aramaic. In Yağmur Heffron, Adam Stone & Martin Worthington (eds.), *At the dawn of history. Ancient Near Eastern studies in honour of J.N. Postgate*, 255–269. Winona Lake: Eisenbrauns.
- Krotkoff, Georg. 1985. Studies in Neo-Aramaic lexicology. In Ann Kort & Scott Morschauer (eds.), *Biblical and related studies presented to Samuel Ivry*, 123–134. Winona Lake: Eisenbrauns.
- Maclean, Arthur John. 1901. *A dictionary of the dialects of vernacular Syriac as spoken by the eastern Syrians of Kurdistan, northwest Persia, and the plain of Mosul*. Oxford: Clarendon Press.
- Macuch, Rudolf. 1976. *Geschichte der Spät- und Neusyrischen Literatur*. Berlin & Boston: De Gruyter Mouton.

- Mann, Jacob. 1931. *Texts and studies in Jewish history and literature*. The Abraham and Hannah Oppenheim Memorial Publications. Cincinnati: Hebrew Union College Press.
- Mengozzi, Alessandro. 2002. *Israel of Alqosh and Joseph of TelKepe*. Corpus Scriptorum Christianorum Orientalium. Louvain: Peeters.
- Mengozzi, Alessandro & Emanuela Braida. 2011. *Religious poetry in vernacular Syriac from Northern Iraq (17th–20th centuries): An anthology*, vol. v, 627–628. Corpus Scriptorum Christianorum Orientalium. Louvain: Peeters.
- Mole. 2015. The three /r/s of Baṛtəlla. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 110–129. Piscataway, NJ: Gorgias Press.
- Murre-van den Berg, Heleen L. 1999. *From a spoken to a written language: The introduction and development of literary Urmia Aramaic in the nineteenth century*. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Mutzafi, Hezy. 2002. Barzani Jewish Neo-Aramaic and its dialects. *Mediterranean Language Review* 14. 41–70.
- Mutzafi, Hezy. 2004a. *The Jewish Neo-Aramaic dialect of Koy Sanjaq (Iraqi Kurdistan)*. Semitica Viva 32. Wiesbaden: Harrassowitz Verlag.
- Mutzafi, Hezy. 2004b. Two Texts in Barzani Jewish Neo-Aramaic. *Bulletin of the School of Oriental and African Studies* 67(1). 1–13.
- Mutzafi, Hezy. 2004c. Features of the verbal system in the Christian Neo-Aramaic dialect of Koy Sanjaq and their areal parallels. *Journal of the American Oriental Society* 124(2). 249–264.
- Mutzafi, Hezy. 2008a. *The Jewish Neo-Aramaic dialect of Betanure (province of Dihok)*. Semitica Viva 43. Wiesbaden: Harrassowitz.
- Mutzafi, Hezy. 2008b. Trans-Zab Jewish Neo-Aramaic. *Bulletin of the School of Oriental and African Studies* 71(3). 409–431.
- Mutzafi, Hezy. 2014. Jewish Neo-Aramaic. *Encyclopedia of the Jews of the Islamic World*. Leiden & Boston: Brill.
- Napiorkowska, Lidia. 2015a. A diachronic perspective on emphasis in Christian Diyana-Zariwaw. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 130–144. Piscataway, NJ: Gorgias Press.
- Napiorkowska, Lidia. 2015b. *A Grammar of the Christian Neo-Aramaic dialect of Diyana-Zariwaw*. Leiden & Boston: Brill.
- Odisho, Edward Y. 1988. *The sound system of Modern Assyrian (Neo-Aramaic)*. Semitica Viva 2. Wiesbaden: Harrassowitz.
- Rockwell, William Walker. 1916. *The pitiful plight of the Assyrian Christians in Persia and Kurdistan described from the reports of eye-witnesses*. New York: American Committee for Armenian and Syrian Relief.
- Sabar, Yona. 1976. *Pašat Wayahī BaŠallaḥ. A Neo-Aramaic Midrash on Beshallah (Exodus). Introduction, Phonetic Transcription, Translation, Notes and Glossary*. Wiesbaden: Harrassowitz.
- Sabar, Yona. 2002. *Jewish Neo-Aramaic Dictionary*. Semitica Viva 28. Wiesbaden: Harrassowitz.
- Soane, Ely Bannister. 1912. *To Mesopotamia and Kurdistan in disguise; with historical notices of the Kurdish tribes and the Chaldeans of Kurdistan*. London: J. Murray.
- Socin, Albert. 1882. *Die neu-aramäischen Dialekte von Urmia bis Mosul*. Tübingen: H. Laupp.

- Talay, Shabo. 2001. Grammatikaliche Anmerkungen und Texte zum neuaramäischen Dialekt von Nerwa (Nordirak). *Mediterranean Language Review* 13. 1–37.
- Tsereteli, Konstantin. 1977. Zur Frage der Klassifikation der neuaramäischen Dialekte. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 127. 244–253.
- Tsereteli, Konstantin. 1978. *The Modern Assyrian language*. Moscow: Nauka.
- Tsereteli, Konstantin. 1990. The velar spirant Ĝ in Modern East Aramaic dialects. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 35–41. Harvard Semitic Studies. Atlanta: Scholars Press.
- Watson, Janet. 1999. The directionality of emphasis spread in Arabic. *Linguistic Inquiry* 30(2). 289–300.

3.5. Iraq-Turkic

Christiane Bulut

1. The Turkman of Iraq

The Turkophone populations of Iraq call themselves Turkman, and their language Turkmanja.¹ The Turkman and their language Turkmanja, or Iraq-Turkic, must be distinguished from the Turkmen language, part of the Northern Oghuz group (see the classification in Table 1 in Bulut, this volume, chapter 4.2, §1), spoken in the Republic of Turkmenistan, in Northeast Iran and in North Afghanistan. Iraq-Turkic is spoken by groups sandwiched between the Kurdish regions to the north and east and Arabic-speaking areas to the south. The speech communities are situated in a number of separate areas, villages and towns within a belt stretching from Tal‘afer in the northwest to Ba‘adra in the southeast (see Fig. 1).² Most larger cities, such as Kirkuk and Arbil, have a mixed, mainly Turkic-Kurdish population. An important group of Turkman is also found in Baghdad. The migration in recent years, which was caused by the Arabization policy of the Baath-Party and subsequent wars and civil wars, has considerably changed the ethnic map of Iraq. Moreover, many villages and traditional environments of the Turkman, such as the old quarters within the citadel of Kirkuk, have been destroyed in recent years.

Established in the aftermath of WWI, Iraq is one of the new nation states of the Middle East whose borders cut through the traditional areas of settlement of various ethnic groups, its multi-ethnic and multi-lingual character partly also reflecting its Ottoman or Iranian heritage. The Turkman of Iraq form the third largest community in Iraq, after Arabs and Kurds; yet, they have been denied the official status of a minority. Due to the repressive policy against the minorities that commenced with the foundation of the new nation state Iraq and was enforced under the regime of the Baath Party after 1968, it is difficult to give precise population figures. In the late 1970s Buluç (1980) estimated that about 750,000 Turkman were living in Iraq. Based on various statistics from pre-colonial data and the last official census conducted in Iraq in 1957, Fischer (1993) arrived at an estimated number of 600,000 for the year 1989, which then amounted to 3.3 % of the total population of Iraq (18.27 million in 1989). This would imply that of today's total

¹ Like other Turkic-speaking communities of the larger region, the Turks of Iraq previously used the ethnonym *Turk*, while referring to their language as *Turki*; in 1959 the military junta introduced the names *Turkman* and *Turkmanja*, which nowadays are widely accepted among representatives of the speech community.

² This Turkman belt is sometimes called “Turkman Eli”.

population of 31.13 million, roughly one million and thirty thousand belong to the Turkman minority.³

Nowadays Iraq is a predominantly Muslim country, with about 65 % Shiites and 32 % Sunnis, and one may assume roughly the same distribution of confessions among the Turkmans. After the collapse of the Baath regime, though, a number of religious minorities, Sufi orders or sects began to re-surface, one of which is the Bektashiyye. Recent sources claim that this Shii-Alevi Sufi order has between 250,000 to 700,000 followers in the Turkman Belt, with centers in the cities of Mosul, Tal‘afer, Kirkuk, Tuzkhurmatı and Arbil.⁴

1.1. Historical background

Iraq displays many features characteristic of a transitional area, a fact that also shaped the prevailing linguistic situation. After the delimitation of the sphere of influence of the Ottoman and the Safavids during the 16th century, major exchanges of populations still occurred in Mesopotamia, which remained part of a disputed border region. Whenever they could gain the upper hand, both Ottomans and Safavids had reliable tribes moved into the area. Shah Ismail relocated a considerable number of Shii tribesmen from the region of Maraghe in Azerbaijan to Iraq. Under Sultan Murad IV (1623–1640) and other Ottoman rulers, Anatolian Turks were settled in this region.⁵ Strategic settlement continued during the 18th century, when Nader Shah (1733–1743) had Shii Turkic tribes brought into the region. At the same time, the Ottomans tried to fortify the Turkic stronghold that lay behind the notoriously unreliable Kurdish principalities by relocating sympathetic settlers from Anatolia in the area. Much of the peculiar crescent-shaped form of the Turkman belt coincides with the environment of the Great Trunk Road from Mosul to Baghdad, which implies strategic movements of local populations in Ottoman times. Toponyms such as Shahseven (i. e. the confederation which formed the military backbone of the Qajar state in the 18th and 19th centuries) indicate that, in later times as well, Turkic tribes adhering to the Shii denomination settled in Iraq.

For the Turkic dialects of Iraq these movements of populations indicate that there was constant influence both from the West and the East, which continued after the formation of the most important written Turkic languages of the region, Ottoman Turkish and Azeri Turkic. It also explains the fact that the map of

³ Recent sources from Azerbaijan and Turkey mention much higher figures; Pashayev (2003: 25) claims that there are about 2.5 to 2.6 million Turkmans living in Iraq, while Turkish media even talk about 3 to 3.5 million. According to a more recent source, the population figure must be at least 2,500,000 (Saatçi 2018: 332).

⁴ Source: the Turkish newspaper ZAMAN, 21 August, 2011.

⁵ See Marufoğlu (1998: 57).

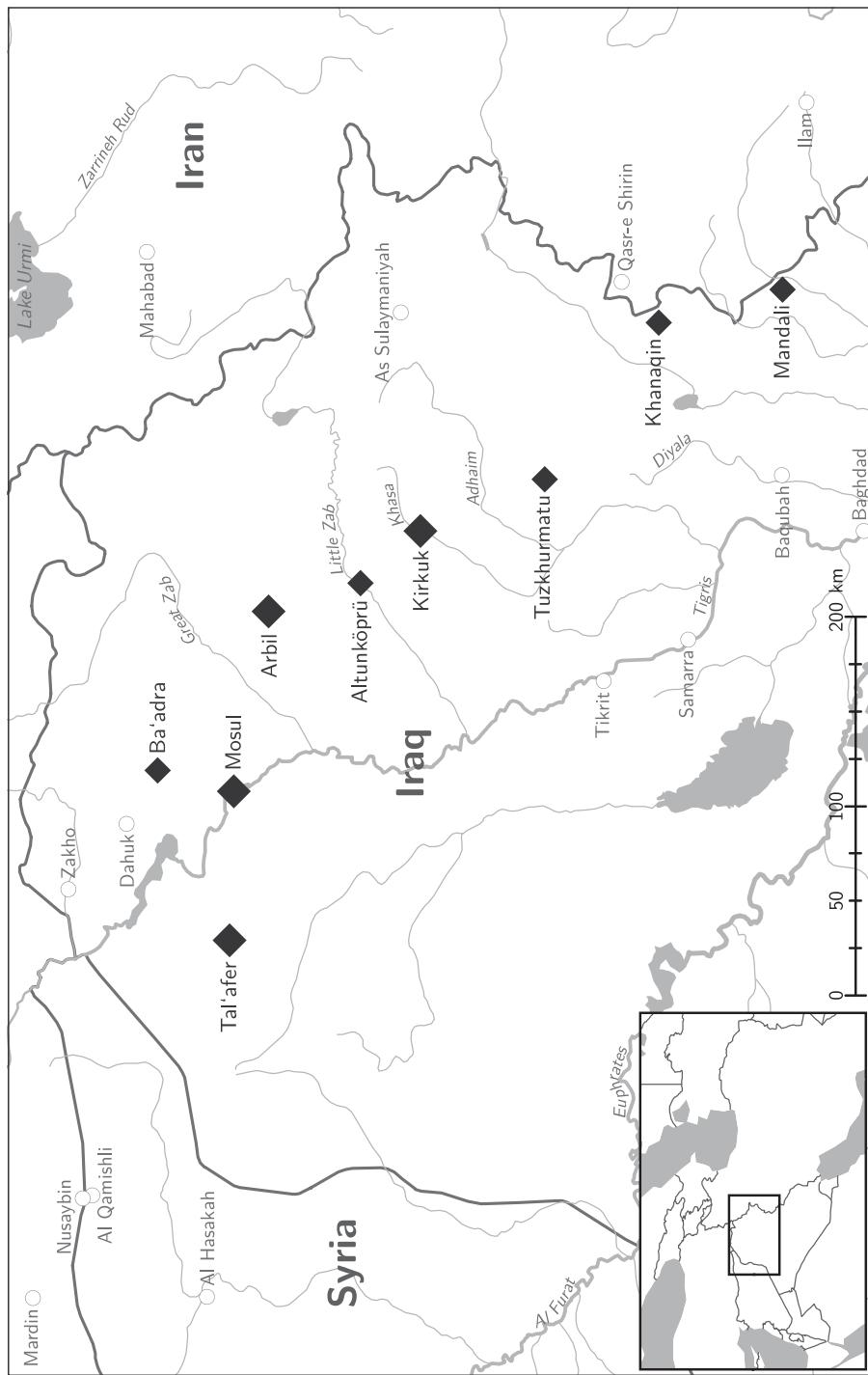


Figure 1: Locations of the Iraq-Turkic varieties in this chapter

Iraq-Turkic dialects is a patchwork rather than a continuum—which is, of course, also true for Anatolian and Iran-Turkic dialects.

1.2. Sociolinguistic situation

Three major languages are spoken in the so-called Turkman belt: Iraq-Arabic, till recently the official language for all of Iraq; Kurdish, which due to the developments on the political scene has gained increasing importance during the last two decades and consequently became the second official language in June 2004; and Iraq-Turkic, which is not used for official purposes. The prestige the latter enjoys is relatively low;⁶ thus, it is rather unusual for native speakers of Arabic or Kurdish to learn Turkic. Most speakers of Iraq-Turkic, who are at the bottom of this hierarchy, are tri-lingual in Arabic, Kurdish and Turkic. Due to a cultural orientation towards Turkey, many also have active command of Turkish of Turkey; additionally, the influence of Turkish media (especially programmes broadcast on satellite TV) has steadily grown during the last twenty years, and a number of private schools, with teaching in Turkish and backed by Turkish institutions, also operate in the region. Thus, diglossia in Iraq-Turkic and Turkish of Turkey is a widespread phenomenon. With the displacement of parts of the Iraqlı Turkman population during the Gulf Wars and the civil war in Iraq, different constellations of intra-Turkman diglossia have also arisen. Nowadays, many speakers of Turkman speak a mixed idiolect displaying various regional features. Despite its low prestige, Iraq-Turkic is also used as a regional language of communication by smaller minorities, such as the Armenians around Mosul and Neo-Aramaic speaking Christians. Iraq-Turkic never developed a written standard of its own. The works on history or folklore, or literary texts comprising short stories and very rarely even novels which appeared in Iraq and in Turkey during the last fifty years are written in Standard Turkish—either in Arabic, Kurdicized Arabic or Latin script.⁷ In 1997, the Iraqi Turkman Congress adopted a Declaration of Principles, Article Three of which states that: ‘The official written language of the Turkman is Istanbul Turkish, and its alphabet is the new Latin alphabet.’⁸

Narrative texts representing the spoken language are rare: Two older collections of folklore texts, the doctoral theses by Choban Khıdır Haydar (1979) and Hussin Shahbaz Hassan (1979), have remained unpublished. Buluç (1973/74), who also wrote some articles on the local varieties of Mandali (1975) and Khanaqin (1979),

⁶ Yet Iraq-Turkic ranks higher in prestige than smaller minority languages such as Armenian, Gorani, and Neo-Aramaic (see Haig, this volume, chapter 3.3, and Khan, this volume, chapter 3.4, for the latter two).

⁷ For a survey on modern Turkman literature see al-Bayati (1970).

⁸ Türkmeneli İşbirliği ve Kültür Vakfı. “Declaration of Principles of the Turkman Congress”. <http://www.iraqiturkman.org.tr/turkmen27.html> (accessed 25 November 2011).

has published a small sample text of Iraq-Turkic from the Tal‘afer region. Different written versions exist of the most famous folklore text, the romantic novel *Arzı ile Qamber*, which the Kirkuk lawyer Ata Terzibashı recorded in the 1960s.

2. Linguistic features

Iraq-Turkic, or Iraqi Turkman are collective terms for a number of transitional dialects and regional varieties of predominantly West-, South- or Central-Oghuz character, which the local populations nowadays call Turkmanja. Being closely related to Iran-Turkic or South-Azerbaijanian, Iraq-Turkic also shares a considerable amount of features with Anatolian-Turkish; consequently, characteristic isoglosses demonstrate affiliations both with the eastern and western neighbours.

The following description will present characteristics of Iraq-Turkic classified according to the linguistic subsystems of phonology, morphology, syntax and lexicon. With regard to their supposed origin, items pertaining to these subsystems will be categorized as (i) genuine Turkic inventory; as (ii) so-called “areal features”, indicating phenomena which appear throughout genetically unrelated languages in a given area; or as (iii) language contact-induced developments which reflect the interaction between structurally different languages. Moreover, cross-references indicate the similarities between the spoken varieties of Turkic in the historical area of East Anatolia, West Iran and Iraq.

Two short sample texts are provided at the end of this chapter, which illustrate some of the linguistic items in context; Text 1 was recorded forty years ago with an old gentleman in Tuzkhurmatı, a predominantly Shiī settlement 55 miles south of Kirkuk.⁹ The informant who provided Text 2 was born in 1961 in Yengi Tis‘ın Mahallesi in Kirkuk; he holds a university degree, and is fluent in Iraqi Turkman, Turkish, Arabic and Sorani Kurdish. In the following, quotations from these sample texts are referenced by a number indicating the text and number of the respective sentence. When referring to other recordings, I will just indicate the place of origin. In addition, the folklore materials mentioned in Section 1.2 were consulted; examples from these texts are quoted with the abbreviation (CKH 1979: number of page/line) for Choban Khidir Haydar (1979) and (HSH 1979: number of text/line) for Hussin Shahbaz Hassan (1979).

⁹ I am grateful to Suphi Saatçi, who shared some of his older recordings of Iraqi Turkman with me.

2.1 Phonology

2.1.1. Consonant inventory

With regard to morphology and lexicon, Iraqi Turkman dialects are quite similar to each other; yet, they differ considerably with regard to phonology. The most important differentiating items are the realization of the old velar /ŋ/ (an Intra-Turkic isophone) and the so-called ‘fronting’ in the articulation of certain stops and affricates (a phonological isogloss in the region).¹⁰

East-West division of Iraq-Turkic dialects – realisation of velar /ŋ/ > [y] or /ŋ/ > [w] as an intra-Turkic isophone

Vertically the Iraqi Turkman varieties can be divided into two groups, using the realization of the old velar /ŋ/ in the possessive suffixes and the possessive type of the conjugation of the 2nd person singular and plural as a morpho-phonological indicator. Around Arbil, and basically also in Kirkuk, the old velar nasal /ŋ/ has developed into [w] or [u]; these varieties belong to the so-called w-group. Most dialects in the west of the Turkman belt, on the other hand, display a development of velar /ŋ/ > [j], representing the y-group. Thus, the pronunciation of the possessive of the 2nd person singular is *seniy* ‘yours’, and *qiziy* ‘your daughter’ in dialects of the y-group, and *senuw* ‘yours’ *qizuw* ‘your daughter’ in the w-group. Due to the mixed background of the population in bigger cities such as Kirkuk, both types of pronunciation may also exist parallel. The speaker from Kirkuk who produced Text 2, for instance, uses forms pertaining to the w-group, such as [gesdou] ‘you (2SG) went through’, and [tʰutsau] ‘if you (2SG) hold’. Tuzkhurmati belongs to the y-group, as forms such as [alardij] ‘you should take’ (1/10) or [gønderdij] ‘you have sent’ demonstrate.

In Iraq-Turkic, some lexical items with irregular distribution also display a development of old velar nasal /ŋ/ into [w] or [u], see examples such as: *soyra*, *sõra* ‘later’ <**soyira*, *gewil* ‘heart’ <**göyül*, *yalyuz* ‘alone, lonely’ <**yalujus*, and [øgynde] ‘in front of’ (2/9) <*ög* ‘front’ <**öŋ*. Across the varieties of the area some few items have consequently preserved the velar /ŋ/ as [ŋ] or even [ŋg], see, for instance, pronunciations of ‘wool’ as [yŋ], [yŋg], [jyŋ], [iŋ].

¹⁰ Stilo (1994) describes this phonological development for Azerbaijan and Transcaucasia; see also Haig, this volume, chapter 3.3, and Khan, this volume, chapter 3.4 for similar phenomena in Kurdish and Neo-Aramaic.

North-South division of Iraq-Turkic dialects - 'Fronting' in the articulation of stops and affricates as an areal isophone

In this process, the velar stops /k/ and /g/ are palatalized to [k^j] and [g^j], or, in the extreme case, affricated to [dʒ] or [f], while the original affricates /dʒ/ and /tʃ/ may be fronted to [dz] and [ts/tç] etc. Thirdly, the place of the former velar stops /k/ and /g/ may be taken by the post-velar or uvular stops, which also shift farther to the front; thus /q/ > [k] and /g/ > [g] also occur.¹¹

Fronting is more predominant in the north-eastern dialects around Arbil, and less explicit towards the South. In the city of Kirkuk, where populations of different origins mix, it is present to a lesser extent. The sample Text 2 displays a relatively strong tendency to fronting, which is characteristic of the neighborhood of Yengi Tis‘în. While /g/ is mostly palatalized (/g/ > [g, g^j, j])—but, to my experience, never affricated to [dʒ]—, /k/ displays palatalized (/k/ > [k^j]) and additional affricated allophones (/k/ > [k^f, ç]) as in [k^eerk^jyç] ‘Kirkuk’ (2/1), or [k^eerem] ‘instance, time’ (2/6). Consequently the unvoiced affricate /č/ has a fronted allophone [ts] (/č/ > [ts]), as in [tsoy] ‘very, much’ (2/1), ['oxutsi] ‘singer’ (2/6), or [getser] ‘it passes’ (2/7). The voiced affricate /j/ is pronounced [dz] in most instances (/j/ > [dz]), see examples such as [gedze] ‘night’ (2/2). Counterexamples, such as [iʃen] ‘drinking’ (2/4), and [eglendʒe] ‘entertainment’ (2/3) obviously reflect the omnipresent influence of Turkish; note that the same speaker may also produce fronted variants in connection with the same lexical items, as, for instance, in [æglendza] (2/7).

The text from Tuzkhurmatı, on the other hand, does not display fronting—nor do the adjacent Iran Turkic varieties. Thus, Kirkuk seems to establish the southernmost point of the fronting area. Note also that fronting is a so-called areal phenomenon which occurs in the pronunciation of genetically unrelated languages in North Iraq, SE-Anatolia and West Iran (see Haig, this volume, chapter 3.3, §3.1.1). In varieties of Arabic spoken in Iraq, the word for ‘lawyer’ *vakil*, for instance, is sometimes pronounced as [wætçil], the Kurdish numerals [tʃphvør] ‘four’ and [pendʒ] ‘five’ appear as [tsp:r], and [pendz] etc. (see Haig, this volume, chapter 3.3, §3.1.1 on fronting in Iraqi varieties of Kurdish).

Note that back [q] and [g] do not take part in the fronting process, see examples such as [qqaja] ‘rock’ (1/2), or [qæræ] ‘black’. The velar fricative *gayn* [γ], which in Turkic appears in word-internal/-final position in combination with velar vowels, is preserved. Word-final /-k/ in some instances becomes [-x]: *yemax* ‘food’, but it also has the voiced allophone [γ], in forms such as [jemay]. Unvoiced stops in

¹¹ As the phonology of Iraq-Turkic is heavily influenced by the surrounding languages where /q/, /k/, /γ/ and /g/ are different phonemes, one cannot interpret these elements as allophones of the Turkic phoneme /k/ (which they would be according to the Turkological tradition).

word final position are sometimes pronounced as unaspirated, see [bøjy^g] < *böyük* ‘large’.

2.1.1.1. Copied elements

Pronunciation of 'ayn [ʕ], bilabial [w] and the pharyngeal fricative [h]

Characteristic dialect markers of Iraqi Turkman and Kurdish varieties from Iraq are copied or contact-induced pronunciations of Semitic phonemes, such as the retracted tongue root glottal stop ‘ayn [ʕ], the pharyngeal fricative [h], the bilabial [w], and the uvular plosive [q], see examples such as [saʕad] ‘time’ (2/7) or [jaʕne] ‘that is’ (2/11), [sajjaħ] ‘traveler’ (1/5), [ahmæd] nom. prop. (1/8), or [hær ħalda] ‘usually’ (2/4), [waxid] ‘time’ (1/10), and [tæqri:ħan] ‘roughly, approximately’ (2/8). Text 2 demonstrates that most of these items appear in combination with lexical items loaned from Arabic, for which speakers from Kirkuk tend to chose the Arabic pronunciation.

In other varieties of Iraq-Turkic, the distribution of the copied items is not always true to the original; speakers may pronounce ‘ayn instead of a normal glottal stop (*hamza*), while the pharyngeal fricative [h] replaces etymological [ħ]. This also happens in connection with Kurdish or Turkic roots. A special case, though, is the Turkic lexeme *hepsi* ‘all’, which speakers from Kirkuk always pronounce as [ħapsi], with word-initial [ħ-].

As the ‘borrowed’ phonetic items do not establish phonemes, the degree of their realization may vary. In the region around Kirkuk, where the influence of Arabic is stronger, the pharyngeal fricative [h] and the retracted tongue root glottal stop ‘ayn [ʕ] are more clearly pronounced than in the dialects around Arbil in the northern parts of the Turkman belt.

2.1.1.2. Unstable word-initial segments

In Arabic, empty syllable onsets are not tolerated, and there is a similar tendency in Kurdish, leading to the insertion of prothetic h-elements (pharyngeal or glottal) in etymologically vowel-initial words. This rule of the contact languages coincides with a Turkic tendency across the whole area to pronounce an additional [h-] before word-initial vowels. Yet, Turkic items such as *his* ‘soot’ (cf. Turkish: *is*), and *hởrmay* ‘knit’ (Turkish: *örmek*), may have preserved an older [h-].¹² In loanwords like *helbet* ‘naturally’ or *heşye* ‘things, belongings’ (a plural of the Arabic *šay* ‘thing’), on the other hand, the secondary or prothetic nature of [h] is obvious.

Another irregular phonotactic process—which, in varying distribution, also occurs in the dialects of eastern Anatolia and western Iran—is the elision of the

¹² On word-initial h- in Turkic see Doerfer (1981/1982).

word-initial /y-/ before the vowels /i/, /ɪ/ and /ü/.¹³ In words such as *ilan* ‘snake’, *il* ‘year’, (corresponding to Turkish *yılan*, *yıl*), the word-initial semi-vowel [y] and the following central-back unrounded vowel /ɪ/ are replaced by [i-], while *igit* ‘hero’ (corresponding to Turkish *yiğit*) has obviously lost its initial /y-/. Similarly, word-initial /yü-/ may be present as /ü-/, as in *üz* ‘face’ (but *yüz* ‘hundred’), *üng* ‘wool’, *üskek* ‘high’, or *üzük* ‘ring’ (compare Turkish *yüz*, *yün*, *yüksek*, *yüzük*).¹⁴

As word-initial sound groups such as *yi-*, *yi-*, or *yü-* are alien to the Iranian contact languages, the reason for this substitution may again be language contact influence. On the other hand, some dialects have preserved older forms like *yıylamax* ‘to weep’, which elsewhere became *aylamak*. Moreover, prosthetic [y-] also occurs, in forms such as *yê:v* ‘house’ for *ev*, or *yeriş-* ‘to reach’ for *eriş-* etc.

2.1.2. Vowel inventory

2.1.2.1. The Turkic vowel inventory

The normal Turkic vowel system consists of eight (+ 1) vowel phonemes, each of which can appear in the stem of Turkic words. They are symmetrically distributed: two unrounded front vowels /i/ and /e/, and two unrounded back vowels /ɪ/ and /a/. Each unrounded vowel phoneme corresponds to a rounded vowel in the same position, namely /ü/, /ö/, /u/ and /o/. The preservation of the ninth vowel, the so-called closed /ê/, is characteristic of the whole area, e. g. the form [je:v] ‘house’. As a rounded allophone of /a/, [ɒ] mainly appears in the vicinity of labialized consonants such as /b/ and /v/ [v, w], as in words such as [?arwɒd] < *avrat* ‘woman’ (2/10) or [bɒba] ‘father’ (1/8). In most varieties of the area Iraq, West Iran/Azerbaijan and East Anatolia, certain words, such as, for instance, *böyük* [bøjyg] ‘large’ (2/6), and *gözäl* [gøzel] ‘nice, beautiful’ (2/1) have preserved the older vowel /ö/, which has developed into /ü/ in Turkish.

2.1.2.2. Copied items in the vowel inventory

In addition to the Turkic vowel inventory, most Turkic varieties of the area have adopted new vowel qualities after the model of Iranian or Semitic languages. Short /a/ in Arabic and Persian loanwords is represented by front half-open to open [a, æ], as in examples such as [mæθælən] < *matalan* ‘for instance’, [baʃas] < *bażż* ‘some’ (2/2 and 2/3) or [bæs] < *bas* ‘enough, however’.

In Turkic stems [æ] is in some instances an allophone of /e/, see examples such as [dælɪg] ‘hole’ (1/3), but also [dælɪg] (1/6); yet, as in the neighboring

¹³ Most of the lemmata quoted here begin with y- in Old or Middle Turkic.

¹⁴ Word-initial /u-/ corresponding to /yu-/, on the other hand, is very rare; HSH (1979: 438) lists a single example: *ut-* ‘to swallow’, corresponding to Turkish *yut-*.

Iran-Turkic dialects, examples for [æ] representing /a/—even in combination with uvular consonants such as /q/—also appear, see, for instance, [qæræ] <*qara* ‘black’.¹⁵

On the whole, Iranicization of the vowel system and the resulting disturbances of vowel harmony (as the main principle of Turkic phonotactics) are less explicit than in the adjacent dialects of Iran.

2.1.2.3. Delabialization of the Turkic front rounded vowels /ü/ and /ö/ as a contact-induced phenomenon

In the southern varieties of Mandali and Khanaqin, the front rounded vowels /ü/ and /ö/ are substituted by their unrounded counterparts [i] and [e] or [i̯e], see examples such as /ü/ > [i, i̯] in [sit] <*süt* ‘milk’, and /ö/ > [e:, e:] in ['e:z] <*öz* ‘self, own’ or [dek-] <*dök-* ‘to pour’. Delabialization of /ü/ and /ö/ is a widespread phenomenon, which points to the influence of neighbouring Semitic and Iranian languages. Thus, across the Turkman belt, East Anatolia and Iran speakers may also centralize or delabialize /ü/ > [u] ~ [i] and /ö/ > [o] ~ [e] due to the interference of these contact languages.

2.1.3. Phonotactic processes, assimilations

2.1.3.1. Shortening or elision of vowels: an Intra-Turkic phonological process

Shortening or even complete elision of the high/close vowel in the first syllable occurs, if this vowel is surrounded by two of the followings sounds: stops, such as [k^h] and [t^h], which mostly display aspiration as a secondary feature, the unvoiced affricate [tʃ], the alveolar fricatives /s/, /ʃ/ or /z/, or the velar fricative [x]. Compare [tʃxandæ] or [tʃxar] (both (1/1), which represent contracted forms of the verb stem čix- ‘to come out’. Shortening or elision of high vowels in a word-initial syllable occurs across the Turkic varieties from North Azerbaijan, Iraq, and West Iran.¹⁶

¹⁵ Across the area, an assimilation of the vowel quality to the famous ‘nine consonants’ – as in Ottoman Turkish – does not occur. These ‘nine (Arabic) consonants’ are: the pharyngeal fricative /ħ/ and the glottal stop ‘ayn /ʃ/, the pharyngealized stops /d/ and /t/, and the fricatives /s/ and /z/, the uvular plosive /q/ and the fricatives /g/ and /ħ/.

¹⁶ The phenomenon obviously is related to the internal structure of the syllable: unvoiced/fortis consonants, especially when combined with aspiration, are longer than voiced, that is, in most instances, lenis consonants. To avoid overlong syllables, the fortis seem to acquire or borrow some of the length of the vowel, which consequently becomes shorter.

2.1.3.2. Metathesis and rhotacism

Throughout the area, different forms of metathesis, lambdacism or rhotacism appear. Consequently, stems and morphemes containing /l/ or /r/ are most susceptible to various kinds of metathesis or assimilations (see below, Morpho-phonology). In most East Anatolian, Iranian and Iraq-Turkic dialects, the word *topraq* ‘ground, soil’, for instance, is pronounced [torpax] or [torpay], compare [^torpayə] ‘its soil’ in (1/2). Moreover, the consonant group /-lq-/ is pronounced as [x], both in Turkic or copied lemmata, see /-lq- > [x]: in [qax] < *qalq-* ‘get up’, or [xax] < *xalq* ‘people’ (1/4).

2.1.3.3. Regressive assimilation /b/ > [m]

Throughout the Turkic dialects of Iraq and Iran, nasals trigger assimilation of the labial consonant [b-]. This type of assimilation affects the pronoun of the first person singular, which is *ben* in older Turkic, but in most instances [mæn] or [men] in Iraq and Iran. Accordingly, the dative and the accusative of the demonstrative pronoun *bu* may have alternative forms such as [muna] instead of *buna* and [munu] instead of *bunu*. Other examples for this type of assimilation are [mindʒiy] ‘pearl’ < *boncuq* and [min-] ‘to mount (a horse etc.)’ < *bin-*.

2.2. Morpho-phonology: assimilations at morpheme borders

Progressive assimilation is very frequent. As a rule, suffix-initial /l-/ assimilates to stem-final nasal, (/ -nl- / > [-nn-]). The abstract suffix {+lIK}, for instance, appears as [+nIK]; e. g. [hormannıγ]—with additional h-prothesis—for *ormanlıq* ‘a wooded place’. Similarly, /-nd-/ is assimilated to [-nn-]; the ablative morpheme changes from {+DAn} to [+nAn]. Accordingly, the ablative of the demonstrative pronoun is pronounced [munnan] < *bundan*. The locative suffix, however, remains {+DA} across Iraq-Turkic varieties. Regressive assimilation of word-final [-r] to the [l-] of the plural suffix {-lAr} is also characteristic of Turkic dialects of the area. Thus, for instance, *olurlar* ‘they become’, is pronounced [olullar] in most Iranian and Iraq-Turkic dialects.

2.3. Morphology

2.3.1. Case morphology

Case morphology displays the typical southwest Oghuz characteristics that appear across Turkish and Azeri or Iran-Turkic dialects. Genitive, dative and locative {+(n)In, +(y)A, +DA} show no conspicuous features. Like in Iran-Turkic or Azeri, the possessive accusative is {+n}, in contrast to Turkish {+nI}. Assimilation of the ablative has been described above. The instrumental case, which is {+(I)nAn}

in most East Anatolian and Iran-Turkic varieties, has the form {+(y)dan}. In the equative case {+cA} the pronominal [n] does not appear, e. g. in forms such as [ardıdʒa] ‘following/behind him’ (< back:POSS.3SG.EQU, confer the corresponding Turkish form with -n-: *ardınca*); an exception is the equative of the demonstrative pronouns, see [ondza] ‘that much’, Text (2/11).

2.3.2. Verb morphology

2.3.2.1. Focal present and aorist¹⁷

Focal present is functionally similar to a present progressive, while what is traditionally termed “aorist” in Turkology is a form covering iterative, habitual, or potential action. Basically, Iraq-Turkic reflects the so-called Azeri type of present paradigms, which it shares with other Turkic varieties from Iran/Azerbaijan, and some eastern Anatolian dialects: The focal present is based on a morpheme {-Ir}; thus, it differs from the aorist in {-Ar} only with regard to its high/close vowel. In practice, though, many of the positive forms of present and aorist have effectively merged. As a rule, monosyllabic stems ending in vowels display the same forms in the aorist and the present paradigm—with the exception of the 3rd person singular, which may be marked by {-I)r} and an additional element {-i} for the focal present, and simply by {-i} for the aorist. Secondly, stem-final liquids /-r/ or /-l/ tend to merge with the /r/ of (mainly, but not exclusively) the aorist marker, while the distinctive vowel is dropped, see, for instance, forms such as *värräm* (< ? *ver-i-r-äm* or *ver-är-äm*) in sample Text (1/9, 1/11).

Thirdly, the distinctive vowel may not be pronounced due to a general tendency to shorten or omit the vowel of middle syllables. Moreover, individual speakers may use different variants of the same aorist or present form, plus copied paradigms from Turkish. Consequently, it is difficult to predict the correct forms and, in some instances, also to analyze them as either present or aorist.

Text 2, for instance, refers to customs and traditions, and one may assume that it is aorist-based. Characteristically, the text displays a mixture of different forms, namely the regular form of the aorist of the 3rd person singular: *gäč-är* ‘pass’ (2/7), but also a special variant of the aorist of the third person singular (verb stem + {-i}), found in (2/6) *gäl-i* ‘comes’ (≈ *gäl-är*), and (2/9) *yiyış-i* ‘come together’ (≈ *yiyış-ar*). The full form of the third person plural aorist occurs in: *oturullar* (2/5) (< *otur-ur-lar*) ‘sit.PL’, *tsalallar* (2/5) (< *čal-ar-lar*) ‘play.PL (music)’, and *baxallar* (2/10) (< *bax-ar-lar*) ‘watch.PL’; in other variants of the same form, though, the syllable containing the differentiating vowel is missing, confer: *düzällär* (2/4) (< *düzäl-är-lär*) ‘prepare.PL’, *gätillär* (2/4) (< *gätir-är-lär*) ‘bring.PL’, *gällär* (2/7) (< *gäl-är-lär*) ‘come.PL’, and *oturlär* (2/7) (< *otur-ar-lar*) ‘sit.PL’. The third singular

¹⁷ The following examples are provided in a normalized orthography.

qal-ır (2/4) ‘remains’ and the 3rd plural *oturullar* (2/5) (<*otur-ur-lar*) ‘sit.PL’ seem to present forms of the focal present.

The overlaps in the forms, and the fact that the meaning of both present types is not always kept neatly apart (one may find present were one expects aorist, and vice versa), creates the impression that either a focal present is not fully developed, or that the two have merged into one. This may be due to the language contact situation, as the surrounding languages (Iranian, Semitic) just have one morphologically simple present. The negative forms of present and aorist, on the other hand, are morphologically distinct, see, for instance, the negative forms of the first person singular of the focal present *ged{-mıräm}* ‘I don’t go (now)’, vs. the aorist *ged{-mäm}* ‘I never go’.

2.3.2.2. *A non-evidential perfect*

As in the adjacent varieties of Iran and Anatolia, the perfect displays a mixed paradigm based on the old morpheme {-mIš}, and the ‘new’ perfect marker in {-y)Ib}, which goes back to the gerund in {-y)Ib} and the copula {+DIr} <*durur* ‘stand.3SG’.¹⁸ The most frequent paradigm consists of {-mIš} for first persons and {-Ip} for second and third persons, or {-mIš} in the first and second persons and {-Ip} for third persons, while Altunköprü, for instance, displays the new perfect marker in {-Ip} in all persons. In contrast to the Turkish perfect in {-mIš}, the perfect occurring in the area (East Anatolia, West-Iran/Azerbaijan and Iraq) is strictly resultative and has no marked inferential or evidential qualities.

2.3.2.3. *The imperative of the second person singular in {-ginen}*

As in most Turkic languages, the bare verbal stem functions as the imperative of the second person singular. Combinations with an additional suffix {-ginen}, which appear across the dialects of the area under observation, express a somewhat milder or more polite form of request or demand.

2.3.2.4. *Potential and impossibility: {-abil-}, {-abilme-}*

As is characteristic of the area, the potential verb form {-abil-} has the negation {-abilme-}, e. g. *gäl-abil-mä-m* (1/9) ‘I will not be able to come.’

¹⁸ According to Bodrogligli (1968: 30), the complex forms in {-y)IbDIr} may have been inspired by the model of the Persian perfect, which—in contrast to the Turkic zero-morpheme—does apply an explicit morphological element to mark the 3rd person singular on the surface. Thus, the complex form *gälübtür* ‘has come’ in the glossary of the Esfahan anonymous, for instance, would be a structural copy of the Persian *âmade ast* (come:passive participle + copula 3SG).

2.4. Morpho-syntactic elements

2.4.1. The question particle {mI}

In the entire area of eastern Anatolia, western Iran and Iraq, the question particle {mI} does not appear. Questions are marked instead by rising intonation and a lengthening of the last syllable—the so-called drawl.

2.4.2. Non-finite verbal forms in dependent clauses

While most non-finite verbal forms have disappeared in the varieties spoken in Iran, Iraqi Turkman displays a rich inventory of gerunds, verbal nouns and participles—which, of course, has important implications for the structure of dependent clauses.

The inventory of participles and verbal nouns comprises: the present participle in {-y)En}, which, like in Iran and Azeri-Turkic may function as a nomen agentis, see, for instance *işlä-yän* ‘worker’, or *bâlxxtut-an* ‘fisherman’; the aorist participles in {-r/-mEz}, appearing in lexicalized units exclusively; the perfect participle in {-mIš}; the so-called future participle in {-y)EcEK}; and the verbal nouns in {-DIK} and {-mEK}.¹⁹ The participle in {-y)En} and the verbal noun in {-DIK} also form the basis of complex temporal gerunds (see below, chapter on syntax). The most frequent gerund markers are: {-yE}, {-yIp}, {-EnI < -ElI}, {-y)IncE}, {aorist+ -kEn}, and {-dIKčE}.

2.4.3. An additional pronominal paradigm based on *bilä*

As in Iran Turkic, Azeri and certain East Anatolian varieties, the pronominal system consists of the analytic pronouns (SG: *män*, *sän*, *o*; PL: *biz*, *siz*, *olar*), while reflexive and emphatic pronouns are based on the stem *öz* ‘self’ (+POSS + CASE). Moreover, Turkic varieties of Iraq and West Iran display a third pronominal paradigm based on the stem *bilä*, which always combines with possessive and case markers. Thus, *bilä-sin-ä* (< *bilä* + POSS.3SG + DAT), in a given context, would mean ‘to him/her’. The examples (1a-b), taken from sample Text 1, are identical in meaning but differ with regard to the pronominal form: In (1a), the pronominal form consists of the personal pronoun of the second person singular in the dative case (*sänä* = you.DAT). In (1b), the pronominal form consists of the stem *bilä*, {-y} or {-w} for the POSS2SG, and the dative marker.

¹⁹ In the northeastern varieties of Iraq-Turkic, the latter has a unique form in {-may}.

- (1) a. (1/9)
- | | | |
|-------------|---------------|-------------|
| <i>para</i> | <i>värräm</i> | <i>sänä</i> |
| money | give.AOR.1SG | 2SG.DAT |
- b. (1/11)
- | | | |
|-------------|---------------|---------------------------|
| <i>para</i> | <i>värräm</i> | <i>bilä-y-ä</i> |
| money | give.AOR.1SG | <i>bilä</i> -POSS.2SG-DAT |
- ‘I will give you money’

2.4.4. Copied morpho-syntactic units

2.4.4.1. *The specifier in {+AkA}*

The copying of bound morphemes is generally taken as an indicator for far-reaching processes of language contact-induced change. One such loaned element is the specifier in {+AkA}, which combines with noun stems and means ‘the aforesaid, the one already mentioned in the context’ and may also be used to mark the head of a restrictive relative clauses, see example (2) below. {+AkA} is a global copy of the Kurdish or, originally, Gorani definite article (cf. Haig, this volume, chapter 3.3), from where it must have spread as far as Sonqor Turkic and the Turkic varieties of the Îl-e Qashqâ’î in Southwest Iran (see Bulut, this volume, chapter 4.2). The copied element is well integrated into Turkic noun morphology. It is attached directly to the stem, precedes the case suffixes, and is even subject to palatal harmony:

- (2) (HSH 1979: 10/13)
- | | | | | |
|----------|-----------------|-----------|-----------------|----------------------------|
| <i>o</i> | <i>oylanakâ</i> | <i>ki</i> | <i>mämläkät</i> | <i>xâxiydi</i> ... |
| that | boy.SPEC | CONJ | home | right.POSS.3SG.COP.PST.3SG |
- ‘The boy whose territorial right it was ...’ (cf. Turkish: *memleket hakkı olan oylan* ...)

2.4.4.2. *Global copies of Iranian enclitic personal pronouns*

In spoken varieties of Modern Persian and in Southern Kurdish, clitic pronouns are extremely frequent. Consequently the Turkic varieties that are in intensive contact with these languages, such as, for instance, southern varieties of Iraqi Turkmen and in Sonqori reflect copies of Iranian clitics. These clitics can be attached to the verb after the personal ending, representing the dative, as in *allam-it* ‘I will buy (for) you’ in example (3), or the direct object, in example (4).²⁰

²⁰ See Buluç (1975: 183), on the dialect of Mandali.

- (3) *almæ* *âlläm=it*
 apple buy.OPT=2SG.DAT
 ‘I will buy **you** an apple/apples.’
- (4) *yäyipti=šan*
 eat.PF.3SG=3PL.ACC
 ‘he has eaten **them**’

In a construction appearing in the Turkic varieties from Iraq and Central Iran, Turkic possessive suffixes have replicated structural properties of Iranian pronominal clitics. The Iranian clitic pronouns are frequently used to index a benefactive or recipient, as in the following example from Sorani Kurdish: *hata mumkîn=mân abê* ‘if possible=1PL may.be’, i. e. ‘if it may be possible **for us**’, where the clitic first person plural pronoun indicates a benefactive (MacKenzie 1961). Especially in combination with the borrowed Arabic noun *lâzim* ‘necessary’ (which are applied to paraphrase the verb ‘to need’), Turkic possessive suffixes appear in a position of equivalence to such Iranian pronominal clitics, and denote the needer/wanter. Thus the Turkic phrases such as *lâzim-im (dir)* ‘I need’ in example (5) below, or *lazim-miz* ‘we need’ appear to replicate Central Kurdish²¹ for ‘I want’, *garak=(i)m-a* (‘necessary=1SG is’), see MacKenzie (1961: 105), where (interestingly) the clitic combines with a copy of the Turkic adjective *gerek* ‘necessary’.

- (5) (HSH 1979: 10/72)
källäsi *lâzim-im* *di.*
 head.POSS.3SG necessary-POSS.1SG COP.3SG
 I need his head. (literally: ‘his head is necessary for me’)

2.4.4.3. {-iš}

An element {-iš} can be attached to finite verbs in the Mandali variety. Buluç (1972=1975: 182) gives various examples for combinations of {-iš} with verb forms in the imperative, the aorist, the past in {-DI}, and perfect in {-mIš}. The suffix {-iš} denotes an intensive mode of the respective verb forms, in the sense of ‘to do something with certainty’, e. g. *ôxi!* ‘read!’, *ôxiš!* ‘make sure that you read’. In all probability, {-iš} is a copy of the Kurdish clitic {-iš} ‘also, even’ [Kurmanjî: *jî*], which, in Central Kurdish (cf. MacKenzie 1961: 128) may be suffixed to either a nominal or a verbal form.

²¹ Modern Persian, on the other hand, expresses ‘to need’ with a formation based on *lâzem* ‘necessary’ and the verb for ‘to have’, *dâran*; ‘I need’ would thus be translated as *lâzem dâram*, literally ‘I have necessary’.

2.4.4.4. Comparative in {+tar} and superlative {+tarin}

The Kurdish morphemes {+tar} ‘more ...’ and {+tarin} ‘most ...’ may appear in combinations with Turkic nouns, e. g. forms such as *irax-tär* ‘farther’ and *čox-tærin* ‘most’, where the Turkic stem *čox* ‘much, many’ combines with the Iranian superlative morpheme {-tarin}.²²

2.4.5. Copied word order properties

2.4.5.1. The dative/directive in postverbal position

Northern Kurdish has a twofold strategy to mark an object to which the action expressed in the verb is directed. As a morphological marker, this object takes the oblique case; syntactically, it is marked by the postverbal position it occupies within the sentence, cf. the following two constructed examples from Northern Kurdish (see also Haig, this volume, chapter 2.3):

- (6) a. *Ez pere didim te.*
 I money give.PRS.1SG you.OBL
 ‘I give you money.’
- b. *Kur-ê mun diçe mektebê.*
 son-EZ my go.PRS.3SG school.OBL
 ‘My son goes to school.’

Central Kurdish also regularly places goals after the verb, but there is no oblique case in Central Kurdish, hence no case marking of the goal. Often, however, the goal is marked through a preposition, e. g. *wazîr kâyazî bird bō kuř* ‘the vizier took the letter **to the boy**’ (MacKenzie 1962: 54). Colloquial Modern Persian, another potential contact language, may also use the postverbal position to indicate the directive. According to the norms of the standard language, directives should be introduced by the preposition *be* ‘to’, compare: (*Be*) *maktab mîravam* (to school I.go). However, in colloquial Persian the form *mîram maktab* (I.go school) is very widespread. Like Kurdish, Iraq-Turkic applies both case marking (dative) and word order to indicate directive, cf. (1) above. This kind of double coding is redundant: In contrast to Kurdish, where the oblique case both denotes the direct and the indirect object, the indirect object *sänä/biläyä* in example (1) above is explicitly specified by its dative case. Consequently, the additional syntactic constraint—postverbal position of the dative—has to be interpreted as a copy of structural features from the contact languages.

²² I observed these copies of Iranian comparative and superlative morphemes in recent recordings with bi-lingual Turkic-Kurdish speakers from the Arbil region; as they exist alongside with the regular Turkic particles *daha* ‘more’ and *en* ‘most’, it is difficult to judge to what extent they are conventionalized.

The same word order rule applies to verbal complexes paraphrasing ‘to begin to ...’. The finite form of the aspect verb *başla-* ‘to begin’ combines with a lexical verb in the form of a non-finite or “nominalized” verb form (the verbal noun in {-mEK}, glossed as VN1, in distinction to other verbal nouns, VN2), which is marked by case (dative), and, additionally, by its postverbal position, as in (7):

- (7) (HSH 1979: 2/33)
- | | | |
|----------------------------|-------------------|-----------------|
| <i>ħâmsî</i> | <i>başlâdîlar</i> | <i>gûlmâyä.</i> |
| all | begin.PST.3PL | laugh.VN1.DAT |
| ‘They all began to laugh.’ | | |

2.4.5.2. Structural copies of *ezafe* constructions

With regard to word order properties, Turkic genitive-possessive constructions with the structure {head/possessor:GEN—modifier/possessum:POSS} are diametrically opposed to the Iranian or Semitic type, where the possessum precedes the possessor {possessum—possessor}. Some Iranian languages additionally connect possessum and possessor by a binding element, the so-called *ezafe*. Bi-lingual Turkic-Kurdish speakers sometimes combine Turkic lexemes using Iranian structural properties, as in: *âd-e xâstâxâne* (name-EZ hospital), for Turkic *xastaxâne-nin âdi* (hopital.GEN name.POSS) ‘the name of the hospital’.

2.5. Syntax

To express dependent clauses most Turkic languages use non-finite verb forms. As a rule, adverbial clauses are based on gerunds, while nominalized verb forms such as verbal nouns and participles are used to express relative clauses or complement clauses.²³ Turkic varieties under strong influence of Indo-European or Iranian (or, to a lesser extent, also Semitic) languages have abandoned these principles of Turkic syntax. The dependent clause is based on a finite verb form. It may be connected to the main clause by a conjunction, which expresses the semantic relation between main clause and dependent clause.

2.5.1. Dependent clauses following Turkic principles of nominalization

While Iran Turkic varieties have nearly completely restructured their strategies of clause combining according to Iranian models, East Anatolian varieties nowadays display most of the non-finite subordinate verbal forms used in Standard Turkish. The position of Iraq-Turkic lies somewhere in between. Non-finite subordinate forms and replicated finite patterns are both used with roughly equal frequency.

²³ See Johanson (1990: 199–200).

It should be noted, however, that the functional distribution of non-finite subordinate forms appearing in Turkish and in Iraq-Turkic in some cases different. The verbal noun in {-DIK} (glossed with VN2), for instance, displays most of the syntagmatic and functional properties it has in Turkish. As a non-finite subordinate verbal form it constitutes the basis of both absolute and attributive relative clauses. In absolute position {-DIK} may combine with possessive suffixes as agent markers, and additional case suffixes. In combinations with the instrumental {+INSTR}, though, it forms complex gerunds which are characteristic of Iraq. On the other hand, the verbal noun does not appear in Turkic-style complement clauses.

2.5.1.1. Temporal clauses

Besides gerunds in {-yE}, {-yIp}, {-EnI (< -ElI)}, {-yIncE}, {aorist+ -kEn}, and {-dIKčE}, morphologically complex forms based on the present participle in {-yEn} and the verbal nouns in {-DIK} and in {-mEK} function as temporal gerunds.

In combination with the suffix in {+deg},²⁴ the present participle in {-yEn} indicates simultaneousness, as, for instance, in: *Irân’ın içinä gir-än-däg* (Iran-GEN into enter-PART.EQU) ‘As we entered Iran ...’ (HSH 1979: 17/27). Across the greater area of East Anatolia, Iran and Iraq, complex gerunds derive from the so-called present participle in {-yAn}, which combines with the case suffixes of the locative in {+DA}, or the ablative in {+DAn} and the postposition *sonra* ‘later’. The resulting forms {-yAnda} and {-yAnNAn} (*sonra*) express temporal relations of simultaneousness (‘as, when’) or posteriority (‘after’); see, for instance (*gün*) *çix-an-da* (come out-PART-LOC, Text 1/1) ‘when the sun rises’, or *män gäl-än-nän sora* (come:-PART-ABL after) ‘after I came’ (HSH 1979: 18/73). In example (8) below, the aspect of simultaneousness (already contained in *gid-än-däi*) is stressed by an additional expression of time, *halat* ‘situation, stage’, which syntactically functions as the head of a relative clause:

- (8) (Own field notes)

<i>mäktäb-e</i>	<i>gid-än</i>	<i>halæd-da, (...)</i>	<i>bašla-rik</i>	<i>Arabca</i>
school-DAT	go-PART	situation-LOC	begin-AOR.1PL	Arabic
<i>oqu-may-a.</i>				
study.VN1.DAT				

‘When we go to school, we begin to study Arabic.’

The verbal nouns in {-mEK} (VN1) and {-DIK} (VN2) combine with possessive suffixes (as agent marker) and the instrumental case in {+(y)dan (< *ilen)} or the

²⁴ Probably the equative suffix; the postposition *deg/değin* ‘up to’ would require a dative on the head.

postposition *birlä* ‘with’ to form temporal clauses, such as, for instance *čal-mây-i-ydan* (set.in.motion-VN1-POSS3SG-INSTR) ‘as soon as he pressed ...’, *oyna-mây-i-birlä* (play-VN1-POSS3SG with) ‘while he was still playing’, or *ye-dig-i-yden* (eat-VN2-POSS3SG-INSTR) ‘as soon as (they) ate, in example (9) below:

- (9) (HSH 1979: 3/141)

<i>otı</i>	<i>yediğiydän</i>	<i>o</i>	<i>qardäšläri</i>	<i>här</i>
greens.ACC	eat.VN2.POSS.3SG.INSTR	DEM	her.brothers	all
<i>altısı</i>	<i>ceyran</i>	<i>oldular,</i>		
six	gazelle	become.PST.3PL		
'As soon as they ate the greens, all her six brothers became gazelles, ...'				

2.5.1.2. Relative clauses

Iraq-Turkic relative clauses based on non-finite/nominalized verb forms, such as the present participle in {-yEn} and the verbal noun in {-DIK}, resemble those of Turkish. In example (10) below, the verbal noun appears in absolute position, fully marked with possessive suffix encoding the agent and case suffix anchoring it in the matrix clause:

- (10) (CKH 1979: 124/143)

<i>Get</i>	<i>bu</i>	<i>de-digim-i</i>	<i>eläginän!</i>
go.IMP	this	say-VN2.POSS.1SG-ACC	do.IMP
'Go and do what I told you!'			

It may also be part of a genitive-possessive construction with the agent of the subordinate form in the genitive, as in example (11):

- (11) (HSH 1979: 10/50)

<i>bunlarin</i>	<i>dediglări</i>	<i>'aynän</i>	<i>köskiünde</i>	<i>bir</i>
they.GEN	say.VN2-POSS.3PL	just_like	breast.POSS.3SG.LOC	a
<i>hä:ritä</i>	<i>kimin</i>	<i>etti</i>		
map	like	do.PST.3SG	'In his heart he took notes like a map of the things they had said'.	

Example (12) demonstrates that the subordinate form in {-yEn} has preserved its verbal potential to take direct objects:

- (12) (CKH 1979: 37/15)

<i>Bu</i>	<i>Allah'in</i>	<i>rähmätin</i>	<i>istiyäni</i>	<i>belä</i>	<i>taptadi</i> , ...
this	God.GEN	mercy.POSS.3SG.ACC	want.PART.ACC	so	beat_up.
'So much did he beat the one who had asked for God's mercy, ...'					

The subjunctions in {-yEn} and in {-DIK} can also be attributed to a head; {-yEn} signals co-reference of the head and the agent of the non-finite subjunction (or its possessor), see example no. (13):

- (13) (HSH 1979: 18/76)

<i>Adam</i>	<i>var</i>	<i>öz</i>	<i>sulbunnan</i>	<i>änän</i>	<i>evlatta</i>
man	exist	own	seed.POSS.3SG.ABL	derive.PART	offspring.LOC
<i>munca</i>	<i>ħürmät</i>	<i>yôxtu</i>			
this_much	respect	not.exists			

‘There are men who don’t see this much respect in their own offspring.’²⁴

If the head co-refers with the object (Acc, Dat etc.) of the relative clause, a verbal noun in {-DIK} is used, as in (14):

- (14) (HSH 1979: 10/134)

<i>Bä:li,</i>	<i>‘ayni</i>	<i>tâni-diy-uv</i>	<i>adam-am.</i>
Indeed	same	know-VN2.POSS.2SG	man.COP.1SG
‘Indeed, I am just the person you know.’			

Examples (10) to (14) demonstrate that Turkmenja applies Turkic structures rendering relativization (pre-positive relative clauses based on non-finite subordinate forms) that are structurally very close to their Turkish counterparts.

2.5.2. Syntactic patterns imitating Indo-European strategies of clause combining

2.5.2.1. Right-branching relative clauses

Relativization may also follow Indo-European patterns, with post-posed relative clauses based on finite verb forms. Restrictive relative clauses often copy Iranian types; frequently the head is introduced by a demonstrative pronoun, the underlying pattern being {DEM + HEAD (+SPEC) + CONJ}, see examples (2) and (15). Eventually, copied markers of specificity may be attached to the head, such as, for instance, the Kurdish unit {+Aka} in example (2).

- (15) *Yerišdix o tâcir-e ki bizi*
 reach.PST.1PL that merchant-DAT CONJ 1PL.ACC
*göndärmîšdi Kermânsâna.*²⁵
 send.PLUP.3SG Kermanshah.DAT
 ‘We reached again the merchant who had sent us to Kermanshah.’

²⁵ Literally: ‘in whose offspring coming from their (‘his’) own seed is not this much of respect.’

²⁶ My transcription after the original recording; see also HSH 1979: 17/40.

2.5.2.2. Temporal clauses with the conjunctive *ki*

A structurally similar pattern forms temporal clauses. It may consist of a demonstrative pronoun introducing a noun, an optional marker of specificity, and the conjunction *ki*, i.e: {(DEM) + NOUN (+SPEC) + CONJ}. In spoken Persian and Kurdish, this pattern is frequently applied to indicate a temporal relation between a main clause and subordinate clause:

- (16) (HSH 1979: 5/216)

<i>Oylan</i>	<i>ki</i>	<i>qápi</i>	<i>sā:si</i>	<i>eşitti</i>	<i>vérdi</i>
Boy	CONJ	door	sound.POSS.3SG	hear.PST.3SG	give.PST.3SG
<i>fızā:hi.</i>					
cry					

‘As soon as the boy heard the sound of the door, he cried out.’

2.5.2.3. Complement clauses

As a rule, Turkic-style complement clauses are based on non-finite verb forms, which are syntactically the complement of *verba dicendi et sentiendi*, such as, for instance, ‘to say’, ‘to hear’, ‘to see,’ or ‘to know’. Turkmanja, though, uses complement clauses containing a finite verb form, which may be introduced by the loaned conjunction *ki* ‘that’, see examples no. (17), (18) and (19) below:

- (17) (CKH 1979: 1/8)

<i>Eşitti</i>	<i>ki</i>	<i>Sultân Murâd</i>	<i>zuhur_edip</i> ,	...
hear.PST.3SG	CONJ	Sultan Murad	appear.GER	
‘She heard that Sultan Murad had appeared, ...’				

- (18) (CKH 1979: 37/9)

<i>Nä</i>	<i>deyipsiz</i>	<i>seleyin!</i>
What	say.PF.2PL	say.IMP.2PL
‘So tell (us) what you (Pl.) said!’		

- (19) (HSH 1979: 5/89)

<i>Bildi</i>	<i>ki</i>	<i>bu</i>	<i>qızçı</i>	<i>ediri</i>	<i>bu</i>	<i>iši.</i>
learn.PST.3SG	CONJ	this	girl.POSP	do.PRS.3SG	this	task.ACC
‘He (the Jew) understood that he (the young man) does it for the girl’s sake.’						

2.5.2.4. Analytic expressions of modality

The notion of finality (intention or purpose/aim to which the action is directed) can be expressed by non-finite subordinate forms, such as, for instance, a verbal noun in {-mAQ} combined with the postposition *için*, see (20, where it is reduced to *-či*), where the postposition and the verbal noun have undergone fusion:

- (20) (HSH 1979: 12/48)

qä:rri salilla' bunu ö:ldürmä:či
 wife send.PRS.3PL he.ACC kill.VN1.POSP
 ‘They send a woman to kill him.’

More frequently used, though, is a construction imitating Iranian models, where a finite verb in the optative/subjunctive, as in example (21) below, or in the imperative, as in example (22) below, replaces the non-finite subordinate form:

- (21) (1/8)

Gälümäzsiz mäni aparasız o däligä?
 come.AOR.NEG.2PL me bring.OPT.2PL DEM hole.DAT
 ‘Won’t you (2PL) come along to guide me to this cave?’

- (22) (Tuzkhurmatı)

Bu kärvän yixti sirhäät etsin.
 this caravan stop.PST.3SG rest make.IMP.3SG
 ‘The caravan stopped to rest.’

The optative/subjunctive or the imperative are typically used in a dependent verb after impersonal modal expressions, such as *gäräk* or *lâzım* ‘(it is) necessary, one has to’, as in *gäräk äl degmäsin* (necessary hand touch:IMP.NEG.3SG) ‘his hand must not touch’, and after modal verbs, such as *istä-* ‘to want’, or *qoy-* ‘to let’ (see also Bulut 2000).

The possibility/impossibility suffix in {-Ebil-} ‘can, to be able to’, and {-Ebilme-} ‘cannot, not be able to’, which in Turkish attach directly to the main verb, may sometimes be carried by a dummy verb *et-* ‘to make/do’, distinct from the main verb. In this manner, the modal expression is split into two parts, the dummy verb plus modal suffix, e.g. *edäbil-* ‘can, to be able to’, and a lexical verb in the optative/subjunctive or imperative. Thus, instead of a compact form expressing inability, such as, for instance, *yiyla-yabil-mir* (weep-IMPOSS-PRS.3SG), the following is found: *ed-äbilm-iri yiyla-sin* (make-POT.NEG-PRS.3SG weep.IMP-3SG) ‘he cannot weep’.

2.6. Lexicon

2.6.1. Turkic items

Turkic vocabulary displays a mixture of distinctively Azeri and Turkish features, reflecting the transitional situation between Anatolian and Azerbaijani/West Iranian dialects.²⁷ Some lexical items overlap, such as the eastern *dal* ‘back’ and

²⁷ The distribution of isoglosses is irregular; note that many Azeri isoglosses also appear in Anatolian dialects.

its western counterpart *arka*, both of which are found in Iraq. Others have a markedly western flavour, such as the verb *sev-* ‘love’, and *eyi* ‘good, nicely’ against the eastern *yaxçı*. Yet, Iraq-Turkic shares a large proportion of its lexicon with more eastern, that is: Iran-Turkic, varieties: adverbs, such as *bu tay* ‘this side’, *qansär* ‘opposite side’, *harda/harada* ‘where’, *indi* ‘now’, *bitov* ‘all, completely’, *kimi(n)* ‘like, till’; nouns, such as *bulay* ‘spring, well’, *bala* ‘child’, *toy* ‘wedding’, *küräkän* ‘bridegroom’, *nänä* ‘mother’, *ušay* ‘child’, *ämçäk* ‘breast’, and verbs, such as *apar-* ‘bring away’, *axtar-* ‘seek’, *tap-* ‘find’, and *işlä-* ‘work’. Prenominal forms based on the stem *bilä-* are characteristic of the area of Iraq and West Iran.²⁸

In some instances, the same Turkic root appears with different derivation suffixes, such as *suvar-* (like in Iran-Turkic), instead of the western *sula-* ‘irrigate’. Exceptional is the formation of to ‘to speak together’: While Turkish uses a reciprocal form of *kon-* ‘put’, viz. *konus-*, Azeri varieties display the same derivation based on a different stem, *daniş-*; in Iraq-Turkic the basic stem is *sele-/söyle-* ‘to speak’; *seleş-* is a typical Iraq-Turkic isogloss. Due to the influence of Standard Turkish, there are also a number of recent loans of Modern Turkish origin, of which Bayatlı (1996: 407) has presented a survey. In general, such loans are more frequent in the written language.

2.6.2. Foreign items

With regard to nouns and adverbs, the lexicon of Iraq-Turkic displays a relatively high number of loans or calques of the Arabic and Kurdish contact languages; compared to East Anatolian dialects, there is also more Persian vocabulary. Loaned nouns are integrated into the Turkic morpho-syntax by means of Turkic possessive suffixes and/or case morphology, see, for instance, complex forms such as *nizik-i-nda* ‘in its neighborhood’, where the Kurdish loaned noun *nizik* combines with the Turkic possessive 3SG suffix and the locative suffix.

Most compound verbs are a mixture of loaned and Turkic material. They typically consist of a loaned nominal element and a calque of the light verb by its Turkic equivalent, e. g. *ħaz et-* ‘love, like’, *xılas elä-* ‘to finish’ and the corresponding Kurdish compound verbs *ħez kırın* or *xılas kırın*. Another strategy for the formation of secondary verbs is the combination with the Turkic denominal verb suffix {+IA}. A great number of verbs originating from such formations are characteristic of Iraqi Turkman, as, for instance, *färäħlä-* ‘to be glad’, derived of the Arabic noun *faraħ* ‘joy, gladness’.

Characteristic idiomatic expressions are found across different languages of the area, e. g. *nänädän ol-* (mother:ABL become) ‘to be born’, cf. Behdinî Kurdish *ji dayik bün* ‘from mother become’, or *yadımdan čıxdı* (remembrance:POSS.1SG. ABL go.out:PST.3SG) ‘I do not remember’.

²⁸ On pronominal systems of the area see Bulut (2003).

Sample texts in a narrow (not strictly phonemic) transcription

Sample text 1: Tuzkhurmatı

- (1/1) iran janinda byr day wør, gyn
iran side.POSS.3SG.LOC INDEF mountain exists sun
tʃxandæ dayin ystynnen tʃxar.
rise.GER mountain.GEN above rise.AOR.3SG
‘There is a mountain close to (the border of) Iran, and when the sun rises it rises above that mountain.’
- (1/2) bu da:yda olan qqjalær bajas qqaja_de:,
this mountain.LOC be.PART rock.PL white rock.COP
a:lti da gø:k tʰorpayı.
down.POSS.3SG too green soil.COP.3SG
‘The rocks that are on this mountain are white, with green soil underneath.’
- (1/3) O dayin ifinde de bir dølig war, kʰimse
that mountain inside also a hole exists nobody
girmez ifine.
go inside.AOR.NEG.3SG inside
‘Inside this mountain is a hole/cave which nobody enters.’
- (1/4) Ne bilsən xax ne war o dølig ifinde?
what know.IMP.3SG people what exists that hole inside
‘So how should people know what is inside this cave?’
- (1/5) b̥yr gyn bir æfendi jælde: d̥uzxurmata
one day a gentleman come.PST.3SG Tuzkhurmatı
dolanırdə, sajjah bir
walk around.IMPF.3SG traveler a
‘One day a gentleman came and walked around Tuzkhurmatı, a traveler,’
- (1/6) æfendi dedi: “borda bu dælig næ:_de?” dedlgr:
gentleman said here this whole what.COP.3SG say.PST.3PL
“wolla ßwɔwa, o qara?ol dayə_de di?eller,
by God, father, that Qaravul mountain.3SG.COP say.AOR.3PL
‘and he said: “What is this hole over there?” They answered: “By God, good man, people say that this is Mount Qaravul,

- (1/7) o da:ýda bir dælig_dyr, qurd
 that mountain.LOC a hole.COP.3SG wolf
 dæligə_dır.” æfendi dede: “Gælmessiz məni
 hole.POSS.3SG-is gentleman said come.NEG.AOR.2PL me
 aparasəz o dælige?”
 bring.OPT.3PL that hole.DAT
 and this is a cave in the mountain, a wolf's hole.” The gentleman said:
 “Won't you (2PL) come along to guide me to this cave?””
- (1/8) kʰimse dʒɔwab vermedə. ßyr gyn gældə
 no_one answer give.NEG.PST.3SG one day came
 dedi: “Gælmæssənəz ahmæd bøba, jæ, sæn
 said come.NEG.COND.2PL Ahmad father yeah you
 məni apar o dælige!”
 me bring.IMP.2SG that hole.DAT
 ‘No one answered. One day he came and said: “If you (2PL) don't come
 along, Ahmad, my man, yeah, you will bring me to this cave!”’
- (1/9) men de dedim øznæ: “wølla məm iʃim
 I also say.PST.1SG himself.DAT by God 1SG.GEN work.POSS.1SG
 war, gæləbilmæ:m!” dedə: “para verrem
 exists come.IMPOSS.AOR.1SG say.PST.3SG money give.AOR.1SG
 sənə.”
 you.DAT
 ‘I said to him: “Honestly, I have work to do, I can't come!” He said: “I will
 give you money.”’
- (1/10) dedim: “e:j, ʃox e:j! para verirse, təz
 say.PST.1SG Eh very good money give.COND.3SG quickly
 alardij ja, o waxıd.”
 take.AOR.PST.2SG yeah that time
 ‘I said (to myself): “Excellent, if he offers you money, you should have
 accepted quickly that time.”’
- (1/11) “næ verim sənə? b'r dæne verrem
 What give.VOL.1SG you.DAT one piece give.AOR.1SG
 bilejə, b'r sari lira, osmanlı lirasə.”
 you.DAT one yellow Lira Ottoman Lira.POSS.3SG
 ‘“And what do I intend to give you? I will give you one, one golden Lira,
 an Ottoman Lira.”’

Sample text 2: Kirkuk

- (2/1) K^erkⁱYçd^e toj tsoy gæzel sejler də, bís toj ...
 kirkuk.LOC wedding very nice thing.PL.COP we wedding
 'There were many beautiful things at a (traditional) wedding in Kirkuk,
 we ...'
- (2/2) Toj halatindæ, mæθ^alæn tojdæn bir
 wedding condition.POSS.3SG.LOC for instance wedding.ABL one
 gedze ?awwal beʃ gyn d' olurði
 night early five day also be.IMPF.3SG
 jadima gjælle:↗
 mind.POSS.1SG.DAT come.PST.3SG
 'When there was a wedding, the first night before the wedding — now I
 remember that it could also be five days (earlier) —'
- (2/3) nedze, evlenmax gynyñnen byr gyn
 however marrying day.POSS.3SG.ABL one day
 ?awwal gedzæse eglendž^e fiklinde
 early night.POSS.3SG entertainment form.POSS.3SG.LOC
 olurða, baʃas.
 be.IMPF.3SG some
 'however, the night one day before the marriage day mostly was
 (celebrated) in the form of a feast.'
- (2/4) mæθ^alæn baʃas ?ajləlere 'qalir iʒgi
 for_instance some family.PL.DAT remain.PRS.3SG alcohol
 iʃen iʒgi dyzeller, tsalixt^ə gⁱætiller (...).
 drink.PART drink prepare.AOR.3PL musician bring.AOR.3PL
 'Usually it is up to the (respective) families, so, for instance, those who are
 drinking alcohol will offer drinks, and they invite musicians.'
- (2/5) otorollar hær halda baxtsada:↗, jaŋ' ha?usdæ, tsalallar.
 sit.PRS.3PL usually garden.LOC that.is yard.LOC play.AOR.3PL
 'Usually they sit in the garden or the yard and play.'

- (2/6) bı' k^čerem 'gæzel 'oxotsilar g^jælɪ', bøjg
 for once good singer.PL come.AOR.3SG great
 oxotsəlariməsdan, ʃabdil wāhid de:, gæzel gæzel
 singer.PL.POSS.1PL.ABL 'Abdu'l-Wâhid also, very good
 oxotsəlariməs war.
 singer.PL.POSS.1PL exists
 'Sometimes there are good singers, some of our great singers, such as
 'Abdu'l-Wâhid²⁹ — we have some really fine musicians.'
- (2/7) gæller otörlar, saʃad bir,ik^ji jædze
 come sit.AOR.3PL time one_or_two night
 gæzel bir æglendza waxid getser (...).
 nicely a entertainment time pass.AOR.3SG
 'People come and sit down, and the entertainment will nicely last till one
 or two at night.'
- (2/8) Bæs indi bu zaman da: tæqri:ʒan bo
 however now this time also roughly this
 'qadet ondza qal'madə.
 custom that.much remain.NEG.PST.3SG
 'However, this tradition has not remained exactly the same.'
- (2/9) Toj gynvnnen byr gyn ?awwæl ge'dzæsə
 wedding day.POSS.3SG.ABL one day early night.POSS.3SG
 byr gæzel qapı øgynde:↗ iʃ'xlar
 one nicely door front.POSS.3SG.LOC light.PL
 janar, aqraba qonʃular jiʃʃe:↗.
 burn.AOR.3SG relatives neighbor.PL come.together.AOR.3SG
 'Now, one lights the lamps in front of the door, and neighbors and friends
 gather there the night before the wedding day.'
- (2/10) Hær_ħalda tsoxlux ērkək qismə, ?arwəd qismə
 usually majority man folk.POSS.3SG woman folk.POSS.3SG
 e:vlərdan baxallar, damlardan baxallar.
 house.PL.ABL watch.AOR.3PL roof.PL.ABL watch.AOR.3PL
 'Usually, it's menfolk (down there), while the women are watching from
 the roofs or the houses.'

²⁹ Probably Abdülvahid Küzecioğlu (born 1925 in Mousallah, a Turkman quarter inside the walls of Kirkuk Castle, died 29th of June 2007).

- (2/11) jañne:, bı̄sde war taxri:bæn ja ondza qis
 that is 1PL.LOC exists approximately yeah that.much girl
 qismə, ?arwod qismə erkeklerden sej
 folk.POSS.3SG woman folk.POSS.3SG man.PL.ABL thing
 edmez, jañni girişmez ondza ojnda
 do.NEG.AOR.3SG that is enter.RECIPR.AOR.3SG that.much dance.LOC
 fejde.
 thing.LOC

'That's roughly what things are like with us: Girls and women's folk don't have business with men, they don't mingle with all this dancing and so on.'

Abbreviations

ABL	Ablative	NEG	Negation
ACC	Accusative	OBL	Oblique
AOR	Aorist	OPT	Optative
COND	Conditional	PART	Participle
CONJ	Conjunction	PF	Perfective
COP	Copula	PL	Plural
DAT	Dative	PLUP	Pluperfective
DEM	Demonstrative	POSP	Postposition
EZ	Ezafe	POSS	Possessive
GEN	Genitive	POT	Potential
GER	Gerund	PST	Past
IMP	Imperative	RECIPR	Reciprocal
IMPF	Imperfective	SG	Singular
IMPOSS	Impossible	SPEC	Specificity
INDEF	Indefinite	VN1	Verbal noun (Infinitive)
INSTR	Instrumental	VN2	Verbal noun (nominalized participle)
LOC	Locative	VOL	Voluntative

References

- al-Bayati, Mehdi. 1970. *Die Anfänge der Prosaliteratur bei den Irak-Türkmenen*. Mainz: Mainz University Dissertation.
- Bayatlı, Hidayet Kemal. 1996. *Irak Türkmen Türkçesi* [The Turkmen Turkic of Iraq]. 664. Ankara: Atatürk Kültür, Dil ve Tarih Yüksek Kurumu Türk Dil Kurumu Yayınları.
- Bodrogliglieti, A. 1968. On the Turkish Vocabulary of the İsfahan Anonymous. *Acta Orient. Hung.* Tomus XXI. 15–43.
- Buluç, Sadettin. 1973/74. Tellafer Türkçesi Üzerine [On Tellafer Turkic]. *Türk Dili Araştırmaları Yıllığı Belleten*, 1973–74, 49–57. Ankara: TDK.

- Buluç, Sadettin. 1975. Mendeli (Irak) Ağzının Özellikleri [The particularities of the Mendeli (Iraq) dialect]. *I. Türk Dili Bilimsel Kurultayına Sunulan Bildiriler 1972*, 181–183. Ankara.
- Buluç, Sadettin. 1979. Hanekin (Irak) Türk Ağzı Üzerine [On the Turkic dialect of Hanekin (Iraq)]. *I. Milleler Arası Türkoloji Kongresi, İstanbul 1973*, 600–603.
- Buluç, Sadettin. 1980. Irak Türk Ağızlarının Bazı Ses Özellikleri [Some phonological particularities of the Iraq Turkic dialects]. *I. Milli Türkoloji Kongresi Tebliğler, İstanbul 1978*. 54/1–54/4. İstanbul: Kervan Yayınları.
- Bulut, Christiane. 1999. Klassifikatorische Merkmale des Iraktürkischen. *Orientalia Suecana*, vol. 48. 5–27.
- Bulut, Christiane. 2000. Optative constructions in Iraqi-Turkman. In Aslı Göksel & Celia Kerslake (eds.), *Studies on Turkish and Turkic Languages. Proceedings of the IXth International Conference on Turkish Linguistics, Lincoln College, Oxford, August 12–14, 1998* (Turcologica 46.), 161–169. Wiesbaden: Harrassowitz.
- Bulut, Christiane. 2003. Pronominal systems in the transitional varieties of eastern Anatolia, Iraq and Western Iran. In Sumry Özsoy (ed.), *Proceedings of the Tenth International Conference on Turkish Linguistics*, 321–335. İstanbul: Boğaziçi Yayınları.
- Bulut, Christiane. 2016. Convergence and variation in the Turkic varieties of Iran: Examples from Qashqâ’i. In Éva Á. Csató, Lars Johanson, András Róna-Tas, & Bo Utas (eds.), *Turks and Iranians: Interactions in Language and History. Proceedings of the Gunnar Jarring Conference, Uppsala 2006*. (Turcologica 105), 235–282. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 1981. Materialien zu türkisch h- (I). *Ural-Altaische Jahrbücher. Neue Folge*. Bd. 1. 93–141.
- Doerfer, Gerhard. 1982. Materialien zu türkisch h- (II). *Ural-Altaische Jahrbücher. Neue Folge*. Bd. 2: 138–168.
- Doerfer, Gerhard. 2006. Irano-Turkic. In Lars Johanson & Christiane Bulut (eds.), *Turkic-Iranian Contact Areas*. (Turcologica 62), 93–113. Wiesbaden: Harrassowitz,
- Fischer, Reinhard. 1993. *Die Türkmenen im Irak*. Berlin: Freie Universität Berlin, unpublished MA Thesis.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A Construction Grammar approach*. Berlin & Boston: De Gruyter Mouton.
- Hassan, Hussin Shahbaz. 1979. *Kerkük Ağzı* [The dialect of Kirkuk]. İstanbul: İstanbul Üniversitesi Edebiyat Fakültesi Türk Dili ve Edebiyat Bölümü dissertation.
- Haydar, Choban Khıdır. 1979. *Irak Türkmen Ağızları*. Metinler – İnceleme – Sözlük [The dialects of Iraqi Turkmen. Texts – Investigation – Dictionary]. İstanbul: İstanbul Üniversitesi Edebiyat Fakültesi Türk Dili ve Edebiyat Bölümü dissertation.
- Johanson, Lars. 1990. Subjektlose Sätze im Türkischen. In Bernt Brendemoen (ed.), *Altaica Osloensis. Proceedings from the 32nd Meeting of the Permanent International Altaistic Conference*, 193–218. Oslo: Universitetsforlaget.
- MacKenzie, David Neil. 1961. *Kurdish Dialect Studies*, vol. I. Oxford: Oxford University Press.
- MacKenzie, David Neil. 1962. *Kurdish Dialect Studies*, vol. 2. Oxford: Oxford University Press.
- Marufoğlu, Sinan. 1998. *Osmanlı Döneminde Kuzey Irak* [North Iraq in the Ottoman period]. İstanbul: Eren Yayınları.
- Pashayev, Gazanfar. 2003. Azerbaijanis in Iraq – Little known people. *Azerbaijan International*. Spring (11.1). 24–25.

- Saatçi, Suphi. 2018. The Turkman of Iraq. In Christiane Bulut (ed.), *Linguistic Minorities in Turkey and Turkic-speaking Minorities of the Periphery*. (Turcologica 111), 331–343. Wiesbaden: Harrassowitz.
- Sümer, Faruk. 1952. Bayatlar. *İstanbul Üniversitesi Edebiyat Fakültesi Türk Dili ve Edebiyatı Dergisi* IV(4). 373–417.

Online resource:

- Türkmeneli İşbirliği ve Kültür Vakfı. “Declaration of Principles of the Turkman Congress”. <http://www.iraqiturkman.org.tr/turkmen27.html> (accessed 25 November 2011).

4. Western Iran

4.1. Western Iran: overview

Geoffrey Khan

1. Preliminary remarks

This section includes chapters on Turkic, Iranian and Semitic languages of western Iran. Some of the languages discussed in these chapters are spoken across a wide geographical area, which extends beyond the confines of western Iran. This applies in particular to chapter 4.2 on Turkic of Iran and chapter 4.6 on Persian. The languages described in these last two chapters extend also into the area covered by section 3 (eastern Anatolia and northwestern Iran). Chapter 4.3 is on Bakhtiari, an Iranian language spoken in the Zagros mountains of southwestern Iran in the four provinces of Khuzestan, Chahar Mahal va Bakhtiari, Lorestan and Esfahan. The description is based for the most part on the documented dialect of Masjed Soleymān. Chapter 4.5 concerns the Iranian language Hawrami, the heartland of which is the Hawrāmān region in the northern Zagros mountains in Kermanshah province. This is a variety of the Iranian language Gorani, the geographical area of which includes northern Iraq (see Haig, this volume, chapter 3.3 on Gorani in Iraq). The description of Hawrami presented here is based on the documented dialects of Pāwa and Nawsūd. Chapter 4.4 concerns the Neo-Aramaic dialects of the region. These include dialects belonging to the subgroups of North-Eastern Neo-Aramaic (NENA) and Neo-Mandaic. The NENA dialects were spoken until the middle of the twentieth century in the Kordestan and Kermanshah provinces and the Neo-Mandaic dialects were spoken in the towns of Ahvāz and Khorramshahr in Khuzestan province. These Neo-Aramaic dialects are now highly endangered and very few speakers are left in western Iran. The descriptions are based on fieldwork carried out in diaspora communities outside Iran. Chapter 4.7 is a description of Kumzari, a language exhibiting a mixture of features from Western Iranian and Semitic spoken in the Musandam governorate of Oman. This profile of Kumzari has arisen through the mixing of migrant communities from Iran and South Arabia.

The languages described in this section exhibit a variety of contact-induced features. These can be classified broadly into features arising through convergence with Persian, the national standard language of Iran, and those that are shared areal features across the languages of western Iran arising through the geographical contiguity of various language communities. Standard Persian has had some impact on all languages of the region, but this appears to be more superficial on

the languages of the areas of Kordestan and Kermanshah provinces (Hawrami and NENA) than on languages spoken in other areas of Iran (Turkic, Bakhtiari, Neo-Mandaic).

Since Turkic and Persian were spoken also in the region of section 3 (eastern Anatolia and northwestern Iran), in the following overview reference will occasionally be made to contact-induced features in languages described in section 3.

2. Phonology

2.1. Zagros d

In several languages of the region a *d* in post-vocalic position undergoes a process of lenition. This has a variety of outcomes across the languages. In the variety of Hawrami described in chapter 4.5 it is realized as an interdental approximant. In several varieties of Hawrami spoken outside Hawraman the weakened *d* is realized as a semivowel /y/. In the speech of migrants from Hawraman in Kermanshah it is realized as /y/ or as a lateral // (Mahmoudveysi and Bailey, chapter 4.5, §3.1). In Bakhtiari *d* after a word-internal vowel or glide is realized as an interdental approximant or in some dialects as a sonorant *r* (Anonby and Taheri-Ardali, chapter 4.3, §2.1). In some Turkic varieties in the Zagros area postvocalic *d* is weakened to the sonorant /r/ or sometimes /y/ in combination with front vowels (Bulut, chapter 4.2, §2.1.2.4). In NENA dialects in the Kordestan and Kermanshah provinces a postvocalic *d* has developed into a lateral // (Khan, chapter 4.4, §5.1.1). The lateral appears both where there is historically a voiced interdental **d* or an unvoiced interdental **t*, e. g. Jewish Sanandaj *'ila* < *'idā 'hand', *bela* < *baytā 'house'. An intermediate stage of development appears to have been **t* > *d*, **d* > *d*, whereby both interdentals became a voiced stop *d*. This intermediate stage is attested in some NENA dialects of northwestern Iran, e. g. Jewish Urmī *ida* < *'idā 'hand', *ade* < *'atē 'he comes'. The lateral // would have, therefore, been the outcome of a lenition of the stop *d*.

2.2. Pharyngeals and pharyngealization

Several languages of the region exhibit the use of pharyngeal consonants. In some languages these are largely restricted to loanwords from Arabic, e. g. Turkic (Bulut, chapter 4.2, §2.1.1.2) and Neo-Mandaic (Khan, chapter 4.4, §5.2.2). In the languages of Kordestan and Kermanshah provinces they are also found in native lexical items. In Hawrami this applies to the unvoiced pharyngeal, which is found in words of Iranian etymology, e. g. *haft* 'seven' (Mahmoudveysi and Bailey, chapter 4.5, §3.1). In the NENA dialects both unvoiced and voiced pharyngeals that are native to Aramaic are preserved in the environment of pharyngealized con-

sonants, or at least those that were pharyngealized at some earlier historical period (Khan, chapter 4.4, §5.2.1). Indeed in both NENA and Neo-Mandaic the original pharyngealized coronal consonants *t* [t^h] and *s* [s^h] now tend to be pronounced without pharyngealization. In the NENA dialects pharyngeals occasionally occur as a reflex of a historical interdental **t* in a pharyngealized environment, e. g. *'ahra* < **'a^hra* < **'atrā* ‘town’.

In the NENA dialects there is a pharyngealized lateral phoneme *l* [l^h], which is stable and kept distinct from its plain counterpart *l* (Khan, chapter 4.4, §5.3.1). The pharyngealized lateral is an areal feature and found also in Hawrami (Mahmoudveysi and Bailey, chapter 4.5, §3.1).

Kumzari has a series of pharyngealized/velarized consonants that have their origin in Semitic, but have spread to words of Iranian etymology, e. g. *pānda* ‘fifteen’. The language has the unvoiced Semitic pharyngeal *h* but not the voiced pharyngeal (van der Wal Anonby, chapter 4.7, §2).

2.3. Laryngeals

In the Turkic varieties of the region a non-etymological laryngeal *h* has developed at the beginning of many words that historically began with a vowel, e. g. *helämi-jälär* ‘they don’t do’ (< *elämiyirlär*), or in loanwords that would have begun with a vowel without the added *h*, e. g. *helbet* (< Arabic *albatte*) ‘naturally’, or *häqiq* (< Persian < Arabic ‘aqiq) ‘agate-stone’ (Bulut, chapter 4.2, §2.1.2.3). This phenomenon appears to be related to the development of the word-initial glottal stop to *h-* in Jewish NENA dialects in the region, e. g. J. Sanandaj *hezəl* ‘he goes’ (< ‘ezəl) (Khan, chapter 4.4, §5.2.1).

In Bakhtiari an intervocalic laryngeal stop develops into *h* in some dialects, e. g. *sāʔat* > *sāhat* ‘hour’ (Anonby and Taheri-Ardali, chapter 4.3, §2.3).

2.4. Realization of *q*

The realization of the postvelar or uvular *q* as a fricative, which is characteristic of Persian (Paul, chapter 4.6, §3.2), has been replicated in a number of other languages in the area. This is found in Turkic varieties of Iran, not only in Persian loanwords but also in Turkic lexical forms, mostly in Sandhi assimilations after voiced consonants or vowels, e. g. *bu qayanası* [bo yejanası] ‘this (was) her mother-in-law’, *oqu-* [oxu-] ‘to read’ (Bulut, chapter 4.2, §2.1.1). In Neo-Mandaic *q* is often realized as a voiced uvular fricative [χ] after a vowel in a stressed syllable (Khan, chapter 4.4, §5.1.2). A similar realization of postvocalic *q* is found in the Jewish Urmi dialect of NENA (Khan 2008: 20). In the Jewish NENA dialects of the region *q* after a vowel or *w*, it is occasionally realized as an unvoiced uvular fricative (Khan, chapter 4.4, §5.1.1), e. g. Jewish Sanandaj *qoqé* [qo:χe:] ‘pots’.

2.5. Palatalization

Turkic varieties in the region exhibit the palatalization of the velar stops *k* and *g* to [k̚] and [g̚], or, in a more advanced form, affrication to [dʒ] or [tʃ], in which case the original affricates [dʒ] and [tʃ] are fronted to [dz] and [ts/tç]. This is found in particular in the Turkic spoken in the north of the region in the Azerbaijani enclave Nakhichevan, in some regions of Georgia, around Tabriz and Urmia in Iran, and in northern Iraq (Bulut, chapter 4.2, §2.1.2.4). Similar palatalization of these consonants is found in the Christian NENA dialects of the Urmia region (chapter 2.5, §5.1). It is also documented in Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §2.1).

2.6. Reduction of high vowels

In the Turkic of the region high vowels are shortened or completely elided in word-initial unstressed syllables, e.g. [tʰke:rde] < *tikärdi* ‘was sewing’ (Bulut, chapter 4.2, §2.1.2.2). A similar reduction of short high vowels can be identified in the neighbouring Jewish NENA dialects, e.g. [ʃi'ma] < *šəmá* ‘heaven’ (Khan, chapter 4.4, §5.5.1), and also in Hawrami, e.g. *wít* ‘he slept’ and *náwt* ‘he did not sleep’ (Mahmoudveysi and Bailey, chapter 4.5, §3.2). In the Northern Kurdish variety of Behdinī, a similar process can be observed in the formation of the present indicative. In most of Northern Kurdish, the present indicative is formed with a prefix [di-], but in North Iraq, the vowel is regularly elided (Haig, chapter 3.3, §4.1).

2.7. Breaking of long vowels

In Bakhtiari long ē sometimes breaks to the sequence *ie*, e.g. *zēne* > *ziene* ‘woman’ (Anonby and Taheri-Ardali, chapter 4.3, §2.1). This process of breaking of the vowel is found also in Neo-Mandaic, no doubt induced by contact, e.g. *biełā* (< **bēłā* < **bayłā*) ‘house’ (Khan, chapter 4.4, §5.4.2).

3. Morphology

3.1. Demonstrative pronouns

In the languages of the region the system of demonstrative pronouns marks two levels of deixis, near and far. This contrasts with the demonstrative systems in dialects of Iranian, Turkic and Semitic in the regions of Anatolia and Iraq, discussed in sections 2 and 3, which generally have more levels of deixis, typically near, far and very far, or near the 1st person, near the 2nd person and far.

There is also some degree of convergence of the form of demonstratives across the languages of the region. This applies, for example, to the form of the demon-

strative pronouns of the Neo-Aramaic dialects (Khan, chapter 4.4, §6.1.1). The NENA singular far deixis demonstrative pronoun *'o* ‘that’ resembles its counterpart in Turkic *o* (Bulut, chapter 4.2, §2.3.2). It is also relevant to note that the gender distinction of this demonstrative has broken down in the NENA dialects. This may also have been induced by contact with Turkic, which has no gender distinctions. The NENA near deixis pronoun *'iya*, which is attested in C. Sanandaj and also in some NENA dialects in Iraq, closely resembles the discontinuous near demonstrative *ī ... a* of Hawrami, e. g. *ī hanār=a* ‘PROX.DEM.ADJ pomegranate=DEM.M’ (‘this pomegranate’) (Mahmoudveysi and Bailey, chapter 4.5, §4.3.5). Compare also the Kumzari near demonstrative *ya* (van der Wal Anonby, chapter 4.7, §4). These are not loanwords in the NENA dialects but rather the result of the shaping of the native morphological material in imitation of corresponding forms in languages in contact.

3.2. The marking of definiteness

The Kurdish definite-marking nominal suffix *-aka* has spread to many languages of the region. In the NENA dialects it has the form *-ăke*, which corresponds to the oblique form of the Kurdish definite article (*-ăkay*) rather than the nominative form (*-ăka*), e. g. Jewish Sanandaj *bela* ‘house’, *belăke* ‘the house’; Christian Sanandaj *besa* ‘house’, *besăke* ‘the house’ (Khan, chapter 4.4, §6.3.1). The suffix is found also in Turkic (Bulut, chapter 4.2, §2.4.2.2), e. g. *śär-ăkă-si-ni* (poem-DEF-POSS.3SG-ACC) ‘this certain/aforesaid poem of his’ (Sonqorî) and Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §3.1.1), e. g. *hast=eke* ‘the bone’. In Hawrami it is inflected for number and gender (Mahmoudveysi and Bailey, chapter 4.5, §4.1.1), e. g. *warg-aka* ‘wolf-DEF.M’ ‘the wolf’, *adā-kē* ‘mother-DEF.F’ ‘the mother’, *palawar-akē* ‘bird-DEF.PL’ ‘the birds.’

Kumzari has a definite-marking suffix with the form *-ō*, e. g. *qisr* ‘palace’, *qisr-ō* ‘the palace’. Unlike in the other languages in this section the definite-marker in Kumzari is attached also to attributive adjectives that modify a definite noun, e. g. *qēşarit-ō jwān-ō* ‘the good date’ (van der Wal Anonby, chapter 4.7, §4, §11).

3.3. The marking of indefiniteness

Indefinite-marking suffixes on nouns are found in many Iranian languages of the region, e. g. Persian *mard-ī* (Paul, chapter 4.6, §4.5.1), Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §3.1.1), e. g. *dovdar=ey* ‘a girl’, Hawrami (Mahmoudveysi and Bailey, chapter 4.5, §4.1.1), e. g. *kuř-ēw* ‘a son’. These constructions may be combined with the cardinal numeral ‘one’ in some of these languages, e. g. colloquial Persian *yek mard-ī* ‘a man’ (Paul, chapter 4.6, §4.5.1) and Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §3.1.1), e. g. *ya dovdar=ey* ‘a girl’. The cardinal may be suffixed in place of the indefinite-marker *-ew* in Hawrami, e. g.

dēq-ya ‘a sorrow’. Kumzari has an indefinite-marking suffix with the form *-ē*, e. g. *qisr* ‘palace’, *qisr-ē* ‘a palace’ (van der Wal Anonby, chapter 4.7, §4).

Turkic (Bulut, chapter 4.2, §2.3.2) and NENA dialects (Khan, chapter 4.4, §6.3.1) of the region use only the cardinal numeral ‘one’ to express indefiniteness.

Neo-Mandaic borrows the Iranian pattern cardinal numeral + noun + indefinite suffix along with the Iranian morphemes *ya* ‘one’ and *-i*, e. g. *ya tājer-i* one merchant-INDF ‘a merchant’ (Khan, chapter 4.4, §6.3.2).

3.4. Copulas

All languages of the region have enclitic present indicative copulas. There is a close, and in some cases complete, correspondence between the inflection of such enclitic copulas and the inflectional endings of present verbs. In some languages the 3rd person forms of the copula, or at least the 3rd singular forms, are distinct from the present verbal inflection, e. g. Hawrami (Mahmoudveysi and Bailey, chapter 4.5, §4.3.5), standard Persian (Paul, chapter 4.6, §6.5), Turkic (Bulut, chapter 4.2, §2.3.3.1), the Christian NENA dialect of Sanandaj (Khan, chapter 4.4, §7.1) and neo-Mandaic (Khan, chapter 4.4, §7.1). There is evidence of diachronic development in some dialectal varieties of these languages resulting in a convergence between the enclitic copula and the present verbal inflection also in the 3rd person. This applies, for example, to colloquial Persian and the Jewish NENA dialects of the region. Persian also has an independent stressed form of copula in addition to an enclitic present copula (Paul, chapter 4.6, §6.5). An independent copula also exists in Neo-Mandaic, which is more frequently used than the enclitic copula. It is obligatorily used when the predicate is a nominal (Khan, chapter 4.4, §7.1). The Neo-Mandaic enclitic copula is restricted to locative or adjectival predicates (Khan, chapter 4.4, §8.1.1).

3.5. Pre-verbal indicative markers

The Iranian languages and Neo-Aramaic of the region share the common feature of using a preverbal particle to mark the indicative form of the present:

Persian:	<i>mī-</i>
Hawrami	<i>mi-, ma-</i>
Bakhtiari	<i>e-</i>
NENA	<i>k-, g-</i>
Neo-Mandaic	<i>qa-, qə-</i>

The Iranian languages also mark the irrealis subjunctive by a preverbal particle before the present verb. There is no particle that corresponds to this in the Neo-Aramaic dialects:

Persian:	<i>be-</i>
Bakhtiari	<i>be-</i>
Hawrami:	<i>bi-</i>
NENA	<i>Ø-</i>
Neo-Mandaic	<i>Ø-</i>

Kumzari expresses the realis by a *d* infix before the inflectional ending. The unrealis is expressed by the absence of this infix (van der Wal Anonby, chapter 4.7, §5.3).

4. Syntax

4.1. Progressive constructions

The pattern of some progressive constructions have spread across linguistic boundaries.

One example of this is the progressive construction formed by the reduplication of the verbal stem reported to exist in Hawrami (Mahmoudveysi and Bailey, chapter 4.5, §4.3.6), e. g. *řamāy ma-ram-ūna* ‘I am now running’. This pattern is found also in the Jewish NENA dialects of the region, e. g. *gāroše gārəš* ‘he is pulling’ (Khan, chapter 4.4, §7.1).

In Persian the progressive is expressed by combining the verb ‘to have’ with the indicative present *dāram mī-ravam* ‘I am going’ (Paul, chapter 4.6, §6.11). Similar constructions are found in some Iranian languages of the Caspian region. In such languages the auxiliary verb related etymologically to the Persian verb ‘to have’ is what Stilo in chapter 5 on the languages of the Caspian region calls a ‘locative be-verb’ (§6.5.8.1). This is evidently its original meaning. So the construction expresses location within an action. The Persian construction has been replicated in some Jewish NENA dialects (Khan, chapter 4.4, §7.1), e. g.

- (1) *hite k-e*
 POSS-3MS IND-COME.PRS.3MS
 ‘he is coming’

It is also relevant here to mention the Bakhtiari indicative preverbal particle *e-*. According to Anonby and Taheri-Ardali this is used with some verbs to express specifically progressive aspect, e. g. *davn-om* ‘I run’ vs. *e-davn-om* ‘I am running’. This *e-* particle is homophonous with the Bakhtiari 3s copula clitic and this may indeed be its origin. If so, this would resemble the progressive construction attested in the Christian NENA dialect of Sanandaj, which is formed by combining a copula with a present verb (Khan, chapter 4.4, §7.1), e. g.

- (2) *garš-en=yen*
 pull.PRS-D.1MS=COP.1MS
 'I am pulling'

4.2. Perfect constructions

The pattern of some constructions of the perfect have spread over linguistic boundaries in the region.

The perfect construction consisting of a participle and copula that is a feature of Persian, e. g. *rafte-am* 'I have gone' (Paul, chapter 4.6, §6.8), Hawrami (Mahmoudveysi and Bailey, chapter 4.5, §4.3.6) and Kurdish dialects of the region, has been replicated in the Jewish NENA dialects *zila=yena* 'I have gone' (Khan, chapter 4.4, §7.1). Also in the Turkic varieties of Iran this pattern of perfect is replicated in constructions consisting of a gerund in {-y}Xb} and copula, replacing earlier Turkic perfect constructions with the morpheme {-mXš} (Bulut, chapter 4.2, §2.3.3.2).

Another pattern of perfect construction consists of a fossilized invariable 3s copula element combined with a past perfect form. This is attested in the Christian NENA dialect of Sanandaj, e. g.

- (3) *g-i-grəš-le*
 IND-COP.3MS-pull.PST-L.3MS
 'he has pulled'

The same construction can also be identified in the perfect of Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §3.3.2), e. g.

- (4) *xavsid-om-e*
 sleep.PST-1S-COP.3S
 'I have slept'

The formation of a present perfect from the combination of a copula and a past verb is attested in Early Judaeo-Persian (Paul, chapter 4.6, §6.10), e. g.

- (5) *nibiš hest*
 write.PST.3S COP.3S
 'he has written'

4.3. The syntax of pronominal clitics

In Persian, and the Iranian languages of the region, the series of pronominal clitics/ suffixes that are attached to nouns to express a possessor is also used to express pronominal direct or indirect objects of verbs, e. g. Persian (Paul, chapter 4.6, §4.9.1):

- (6) a. *barādar=at*
 brother=2s
 ‘your brother’
- b. *mī-bīnam=at*
 IND-see.
 PRS.1S=2S
 ‘I see you’
- c. *gostam=at*
 say.PST.1S=2S
 ‘I told you’

In some Iranian languages of the region such as Persian and Bakhtiari there is generalized accusative alignment and these pronominal clitics express the objects of both present and past verbs. In other Iranian languages, such as Hawrami, where there is ergative alignment, the clitics express the object only on present verbs.

The functional range of the pronominal clitics in Turkic varieties in Iran has been extended from possessive clitics on nouns to clitics expressing benefactive and dative on predicative constructions, e. g. *lazim-miz* (necessary-POSS.1PL (is)) ‘we need’, *ertebāt var-i* (relation existent-POSS.3SG), literally ‘his contact (is) existent’, that is: ‘he has contact’ under the influence of Iranian languages (chapter 5.1, §7.2.3).

Neo-Mandaic, which has accusative alignment with all verbal forms, uses the series of possessor pronominal clitics to express the pronominal objects of past verbs and of 1st and 2nd person present verbs. 3rd person present verbs express the pronominal object by oblique clitics consisting of the oblique, originally dative, preposition *l-* and pronominal suffix (Khan, chapter 4.4, §7.2). In Classical Mandaic the ancestor of the Neo-Mandaic present forms, i. e. active participles, could express pronominal objects with either possessor suffixes or oblique *l*-suffixes also when the subject was 1st or 2nd person (Nöldeke 1875: 391). This suggests that the Iranian-type of pronominal objects have gained ground in the Neo-Mandaic present, most likely through convergence with the Iranian model. The NENA dialects, which have accusative alignment with present verbs, normally use oblique *l*-suffixes to express the accusative pronominal object of present verbs. There is not, therefore, a correspondence between the possessor suffixes on nouns and those expressing verbal objects as in Iranian languages. In the Jewish NENA dialects of western Iran, however, unlike NENA dialects elsewhere, the possessor suffix is used to express the pronominal object of present verbs with 1st person singular subjects (Khan, chapter 4.4, §7.1), e. g.

- (7) Jewish Sanandaj (Khan 2009: 154–155)
gārəš-n-ef
 pull.PRES-1S-POSS.3MS
 ‘I (ms) pull him’

This could also be interpreted as convergence towards an Iranian-type of expression of pronominal object. As in Neo-Mandaic this convergence has taken place in verbs with non-3rd person subjects.

4.4. Oblique case marking of nouns

Hawrami makes a distinction between nominative and oblique case in nouns. The contexts in which the oblique case is used include (i) the complement of an adposition, (ii) the dependent element in a genitive (*ezāfe*) construction, (iii) the argument expressing the object of a present tense verb and (iv) the agent of a past transitive clause (Mahmoudveysi and Bailey, chapter 4.5, §4.1.1). In the Jewish NENA dialects of the region oblique case marking has developed in demonstrative pronouns by bonding the genitive particle *d* with the pronoun, e. g. *do* ‘that (oblique)’ <*d* GEN+*o* DEM (for this process see Khan 2015, 2016, vol. 1: 215–217; Gutman 2018). Such oblique demonstratives are used not only as complements of adpositions and in genitive constructions in noun phrases, which are the historical contexts of the use of the genitive particle *d* in Aramaic, but also occasionally as the direct object of present verbs, e. g.

- (8) J. Sanandaj (Khan 2009: 157)
- | | |
|----------------|--------------|
| <i>do</i> | <i>gārəš</i> |
| DEM.OBL | pull.PRS.3MS |
| 'he pulls him' | |

This extension of the contexts in which the oblique pronoun is used is likely to be the result of convergence with Iranian oblique constructions.

4.5. Postverbal position of goal constituent

A shared feature of the languages of the region is the placement of the constituent expressing the goal of an action after the verb. This applies also to the Turkic of Iran, in which a constituent with dative case expressing a goal occurs in postverbal position. This deviation from the regular verb-final syntax of Turkic languages in other regions must have developed by convergence with the Iranian languages of the area (Bulut, chapter 4.2, §2.4.2.4).

4.6. Alignment

The Iranian language Hawrami has ergative alignment in clauses with past perfective and present perfect verbal forms. Ergativity is found in the same constructions also in the Jewish NENA dialects of Kordestan and Kermanshah provinces, with distinct inflectional endings marking the subject of transitive and intransitive verbs. With the exception of the Jewish dialect of Sulemaniyya and Halabja

across the border in Iraq (Khan 2004), in other NENA dialects the subject of transitive and intransitive past perfective and perfect verbs have the same inflectional markers. It is likely that the ergativity of the Jewish dialects of western Iran arose by convergence with the neighbouring ergative Iranian languages.

Persian and Bakhtiari have accusative alignment. Neo-Mandaic is likewise accusative in its alignment. This is due to the fact that it has been in contact with Iranian languages with accusative alignment, such as Persian and Bakhtiari, rather than ergative Iranian languages.

4.7. Adverbial temporal clauses

Many languages of the region exhibit a common syntactic pattern of adverbial temporal clauses that stand in sentence-initial position. In this pattern the subordinating particle (glossed below as SUB) of the adverbial clause is placed immediately before the verb, which is preceded by one of its complements (typically subject or object), e. g.

- (9) Bakhtiari (Anonby and Taheri-Ardali, chapter 4.3, §4.3.2)

ya modat=ey ke godašt
a while=INDEF SUB pass.PST.3SG
'When a while passed, ...' (= 'after a while ...')

- (10) Turkic (Bulut, chapter 4.2, §2.5.2.2)

o šoxm ke elämišdik äkärdik.
that plough SUB do.PLUPRF.1PL sow.IMP.1PL
'After we had ploughed (the fields), we used to sow.'

- (11) NENA, Christian Urmi (Khan 2016, vol. 2: 490)

+búsra kat-parəm-va, ...
meat SUB-slaughter.PRS.3MS-PST
'When meat was slaughtered, ...'

Similar constructions with the subordinator shift to preverbal position in the languages of the Caspian region are described by Stilo in his chapter in this volume (§8.2.4).

4.8. Evidential

The languages of the region exhibit a certain tendency to extend the use of the perfect to the expression of the evidential. Kumzari has a more extensive grammaticalized system of evidentiality, which classifies the information source as firsthand sensory, non-firsthand reportive, or non-firsthand inferred (van der Wal Anonby, chapter 4.7, §8).

5. Turkic and NENA

Some of the distinctive features of Turkic described in chapter 4.2 have been replicated in NENA dialects of the Urmi region in the West Azerbaijan province in northwestern Iran, but not in the NENA dialects further south in the Kordestan and Kermanshah provinces. These include the palatalization of dorsal consonants, which has already been mentioned in §2.5 above.

Another phonological feature is the development of suprasegmental pharyngealization in the NENA dialects of the Urmi area (chapter 2.5, §5.4), which must have had a Turkic model. It is important to note that vowel harmony is in the process of breaking down in the Turkic of Iran, resulting in phenomena such as the delabialization of front rounded vowels (Bulut, chapter 4.2, §2.1.2.1). The model of the NENA harmonic systems, therefore, is likely to have been the Turkic of earlier periods.

A syntactic feature of the NENA of the Urmi region that has developed due to convergence with Turkic is the frequent use of infinitives and verbal nouns in subordinate adverbial clauses (Khan 2008: 288–290, 2016, vol. 2: 241–247), which replicates the syntax of Turkic gerund constructions. Bulut, however, reports that in the Turkic of Iran today such subordinate gerund constructions have been largely replaced by subordinate constructions with finite verbs under the influence of Persian (Bulut, chapter 4.2, §2.5.1.1). Here again, therefore, it appears that the convergence of the NENA dialects with Turkic must have taken place at an earlier period when gerunds were still widely used in Turkic. The lexicon of the NENA dialects of Urmi has a large proportion of Turkic loanwords (Khan 2008: 383–385, 2016, vol. 3: 1–3). In some cases these words are no longer used in the contemporary Turkic of Iran, but can be traced to Ottoman Turkish (Khan 2016, vol. 3: 2). In addition to lexical, there are some material borrowings of morphology. The Jewish Urmi NENA dialect, for example, has borrowed the Turkic instrumental case ending *-inan* (for this form of the ending in the Turkic of Iran see chapter 4.2, §2.3.1.3) as a phrasal coordinating particle (Khan 2008: 234–235).

A feature relating to verbal aspect is the functional levelling of progressive and non-progressive present in Turkic. Bulut (Bulut, chapter 4.2, §2.3.3.2) describes how in some varieties of Turkic the distinction between the present progressive (“the focal present” in her terminology) and the aorist, i. e. the non-progressive present (habitual, iterative, intention) has been lost. In some varieties this is due to a formal merger of the aspectual verbal affixes and in some varieties the functional range of the progressive form has been extended. This phenomenon in Turkic may have had an impact on the NENA dialects of the Urmi area, in which the originally progressive verbal form has come to be used with a non-progressive function. This is found frequently in the Christian Urmi (Khan 2016, vol. 2: 185–198) and Jewish Urmi (Khan 2008: 274–278) dialects. In the Jewish Salamas dialect this is regularly the case.

References

- Gutman, Ariel. 2018. *Attributive constructions in North-Eastern Neo-Aramaic*. Studies in Diversity Linguistics 15. Berlin: Language Science Press.
- Khan, Geoffrey. 2004. *The Jewish Neo-Aramaic dialect of Sulemaniyya and Halabja*. Studies in Semitic Languages and Linguistics 44. Leiden & Boston: Brill.
- Khan, Geoffrey. 2008. *The Jewish Neo-Aramaic dialect of Urmî*. Gorgias Neo-Aramaic Studies. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2009. *The Jewish Neo-Aramaic dialect of Sanandaj*. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2015. Domains of emphasis, syllable structure and morphological boundaries in the Christian Urmî Dialect of Neo-Aramaic. In Geoffrey Khan & Lidia Napierkowska (eds.), *Neo-Aramaic and its linguistic context*, 145–161. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmî*. 4 vols. Leiden & Boston: Brill.
- Nöldeke, Theodor. 1875. *Mandäische Grammatik*. Halle: Buchhandlung des Waisenhauses.

4.2. The Turkic varieties of Iran

Christiane Bulut

1. Historical background and socio-linguistic situation

For the varieties of Turkic spoken all over western and central Iran (including some pockets in eastern Iran) and their close relatives in Azerbaijan, eastern Anatolia and Iraq no fully developed classification or Turcological model exists. The linguistic map of the region has many blank spots, and a comparative study of the whole area is still a *desideratum* — although a number of older and more recent studies describe the Turkic varieties of individual regions, or inside the borders of the respective nation states.¹ Even terminology is still a matter of debate: Iran-Turkic is called Turkî by most of its speakers, while the Turks of Iraq call themselves Turkman, and their language Turkmanja. Publications from Azerbaijan use expressions such as, for instance, Azeri (dialects) of Iraq, and Azeri (dialects) of Iran, or South Azeri — often with political implications. Azerbaijani, on the other hand, denotes the official language of the Republic of (North-)Azerbaijan. In Turcological literature, closely related dialects in Turkey and Iraq generally are referred to as eastern Anatolian or Iraq-Turkic/-Turkman dialects, respectively.

The Turkic varieties of Azerbaijan, Iran (with the exception of Khalaj and certain dialects of Turkmen spoken in north-eastern Iran near the border of Turkmenistan²), Turkey proper, eastern Anatolia and Iraq represent the western or Saljuqian branch of the Oghuz language group (see Table 1 below). The members of this branch are genetically related and have shared major historical and linguistic developments. During the 11th century, for instance, their speakers followed the Saljuq's migration from Central Asia into Iran, Iraq and Anatolia – thus causing a first Oghuz split; for roughly four hundred years thereafter they formed part of the tribal confederations dominating the greater area of Azerbaijan, eastern Anatolia, Iran and Iraq, where some predecessor of 'Turkî', an Oghuz variety heavily spiced with Iranian elements, was used as a language of communication.

Due to their common historical background, the Turkic varieties spoken in present day Iran are very similar to each other and to the adjacent dialects west

¹ Intensive research on dialectology has been carried out in North Azerbaijan (Republic of Azerbaijan), see Širaliev (1983) on North Azerbaijani dialects, and the Dialect Atlas (İslamov, M. İ., Aghajev, A. G., et. al. 1990). For a comprehensive survey of characteristic features of Iran-Turkic varieties and models for their classification see Doerfer (2006) and a very interesting model developed by Windfuhr (Enc. Iranica 2006).

² Namely Turkmen (called North Oghuz in Doerfer 2006); Khorâsân-Turkic (or, according to Doerfer 2006, East Oghuz) denotes a group of transitional dialects.

of Iran, with characteristic features and isoglosses reaching down into Iraq and eastern Anatolia. The official or standard language to which these Turkic varieties are most closely related is Azerbaijani of the Republic of Azerbaijan, while Turkish of Turkey has developed differently for at least 500 years.

All Turkic varieties of this region display different degrees of Iranicization. In Iran proper, Iranicization of phonology, syntax and lexicon is extremely strong. This facilitates mutual intelligibility despite some minor regional variation.³

The following survey (Table 1) lists the main representatives of the Oghuz language group according to Doerfer's (2006) classification (marked with an asterisk *), which is based on geographical criteria; it also considers the historical affiliation of different Oghuz groups with the Saljuqs' movement to the West and their relation to Turkî, the predecessor of present day Azeri, Iraq and Iran-Turkic. Figure 1 indicates the locations of the groups mentioned in this chapter.

Table 1: The Oghuz language group

Name	*Geographical classification according to Doerfer (2006)	Historical affiliation	Spoken in
Turkmen	*North Oghuz	non-Saljuqian	Republic of Turkmenistan, border regions of NE Iran
Khorâsân Turkic	*East Oghuz	transitional dialects, mainly Saljuqian > Turkî	Khorâsân province in NE Iran
Azerbaijanian Turkic/Azeri	*Central Oghuz	Saljuqian > Turkî	Republic of Azerbaijan, Azerbaijan province in NW Iran; exclaves Galûgâh, Lotfâbâd, Daragaz (Khorâsân); Terekeme in Kars province etc.
Qashqâ'i and related dialects	*South Oghuz or Afshâr	Saljuqian > Turkî	SW Iran (South of Qazvîn/Zanjân: Central province, Sonqor, provinces of Esfahan, Shirâz, Kerman; exclave: Qabul/Afghanistan.)

³ The influence of the mainstream Oghuz varieties is very strong and has even invaded Khalaj, the only non-Oghuz member of the Turkic languages of central Iran: Nowadays Khalaj is a strongly Oghuzicized and Iranicized language containing some relics or fragments of an underlying older Turkic linguistic system.

Name	*Geographical classification according to Doerfer (2006)	Historical affiliation	Spoken in
East Anatolian dialects	transitional area between Azeri and Turkish	Saljuqian > Turkî	SE and NE Anatolian dialects; Urum of Georgia
Iraq Turkic	transitional area between Central, South and West Oghuz	Saljuqian > Turkî	Turkman belt in Iraq
Turkish of Turkey and related dialects	*West Oghuz	Saljuqian > Turkî	Turkey and periphery: Balkans, Gagauz

1.1. Formative factors in the recent history of settlement in western Iran

In the early 16th century, new political and religious powers began to emerge in the Middle East, namely the Ottomans (espousing Sunni Islam) in the West, and the Safavids, their main competitor, implementing the Shi'a of the Twelve Imams as their state religion in Iran. A highly disputed border cut through the area of settlement of the local Turkic speaking populations, which laid the grounds for the diverging development their language — a second major split in the Oghuz language group, into Ottoman and Azeri.

Initially, the Safavid rulers of Iran resided in Tabriz in Azerbaijan (the old capital under the Aq- and Qaraqoyunlu); later the political centre moved to Qazvîn, still in the northwest of the country, and then to central-western Esfahan (already an important centre under the Saljuqs).

Although the borders to the West (the Ottoman-Safavid border) had become more defined, the fluctuation and the dislocation of especially nomadic elements of the population inside Iran was considerable. Two events in the recent history of Iran caused further major changes in the landscape of settlement. Many tribes and communities were displaced during the raids of the Afghan Ghilzay tribes in 1722. In the subsequent restitution period under Nader Shah (1736–1747), the political centre shifted to Khorasan. While Nader Shah attempted to fortify his new stronghold in the Northeast, Turkic and Iranian tribes from all over the country were exposed to deportation or forced settlement, among them Kurds, Turks and Lurs from the Province of Fars. The establishment of the Kurdish Zand dynasty (1750–1779) took place in Shiraz in the extreme Southwest. There, tribes returning from the exile in Khorasan met with other followers of the Zand; among those were also Turkic-speaking communities from the northwest, the area of Azerbaijan proper.

Moreover, most established rulers used to send the tribes all over the country, locating them in sensitive areas at the frontiers, re-locating unruly elements in marginal regions, or forming powerful tribal confederations to back up their political influence.

Thus, the relocation of populations between the 16th and 18th centuries depicted in Fig. 2 are highly characteristic of the history of settlement in Iran. Uprisings of tribes requiring severe measures such as deportation or forced settlement are by no means a phenomenon of the remote past; they still took place during the first decades of the 20th century in the region of Iran and Anatolia.

1.2. Linguistic situation in present day Iran

Nowadays the linguistic situation of Iran is still a multi-faceted mosaic of different languages, including the Turkic, Iranian, Semitic, and Kartvelian language families, with different constellations of bi- and multilingualism prevailing across the area. Consequently, traces of contact-induced linguistic change can be detected throughout the region.

Given the political dominance that Turkic dynasties exerted over the greater region of Iran for roughly nine centuries (the Qājār lasted till 1925), the relatively low esteem in which Turkic languages were held is somewhat surprising. Ever since the arrival of the Saljuq Turks, Modern Persian, the language of the administration and literature, had maintained its place as the most prestigious language. In recent times, it has also gained considerable influence as a means of communication, the language of the media and of education — a development that was accelerated under the Pahlevis, who adopted the concept of linguistic homogeneity as an important means of nation building. For nearly a millennium, Turkic has been spoken widely, although it never reached the status of a prestige or standard language.

Traditionally, Turkic had also developed into an supra-regional language of communication among smaller communities of speakers of minority languages (such as, for instance, the Armenians of eastern Anatolia, Iraq and Iran; or Kurds, Georgians, and Armenians in the Province of Esfahan etc.), and it is used as a means of communication among people of different ethnic backgrounds in certain social and professional groups, such as, for instance, the Bāzāris of Tehran, and, far away from our area of concern, the street vendors in Moscow.

1.2.1. The speakers of Turkî

Most Turkic-speaking inhabitants of Iran adhere to the Shi‘i denomination of Islam (also called Shi‘a of the Twelve Imams, or Ja‘fariyya). They call themselves *Turk* or *Turkî-zabân* ‘turcophone’, and their Turkic language Turkî/Torkî. The expression Azer/Azerî is used only in the northwest, in the Iranian province of Azer-

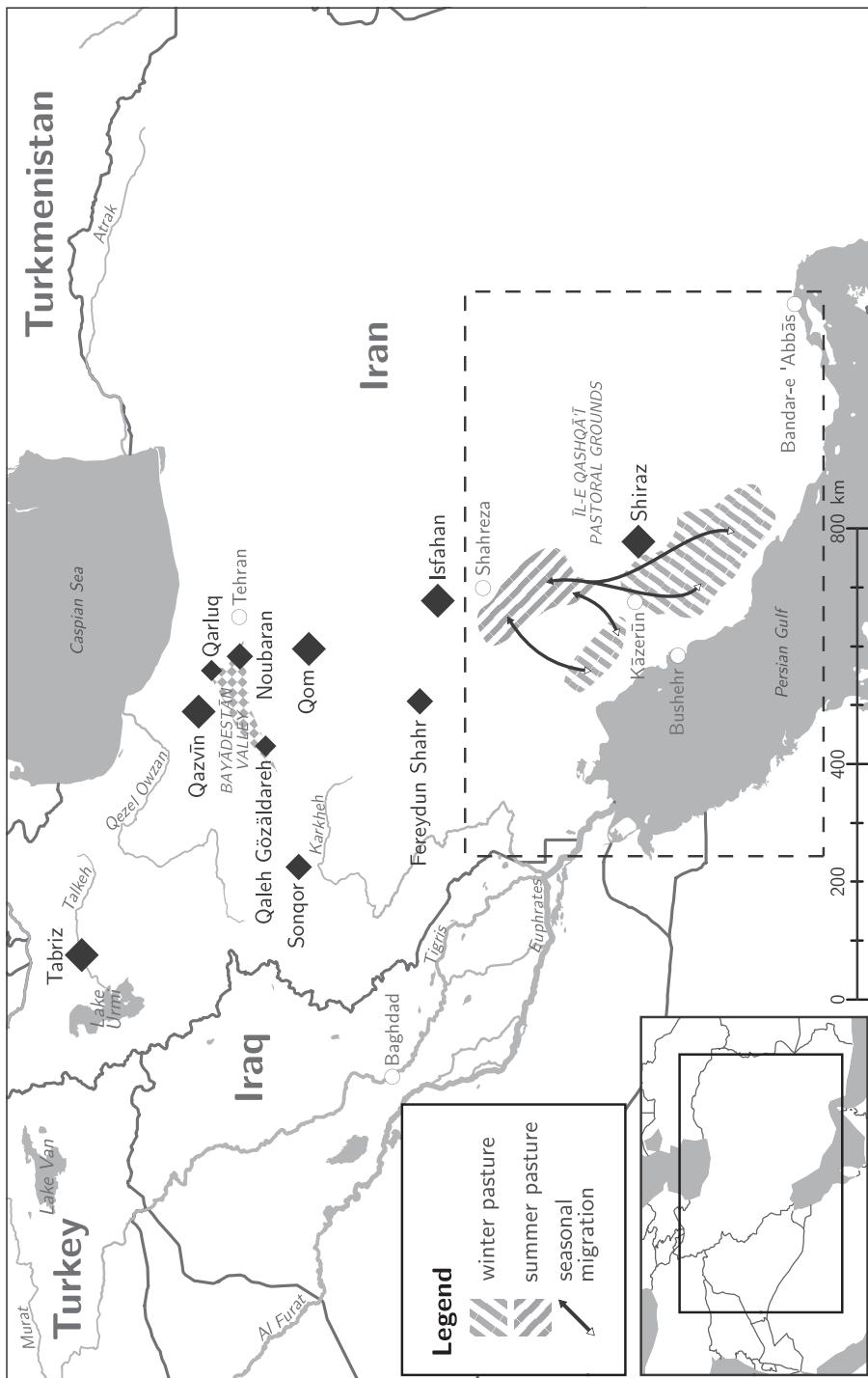


Figure 1: Locations of the Turkic varieties mentioned in this chapter

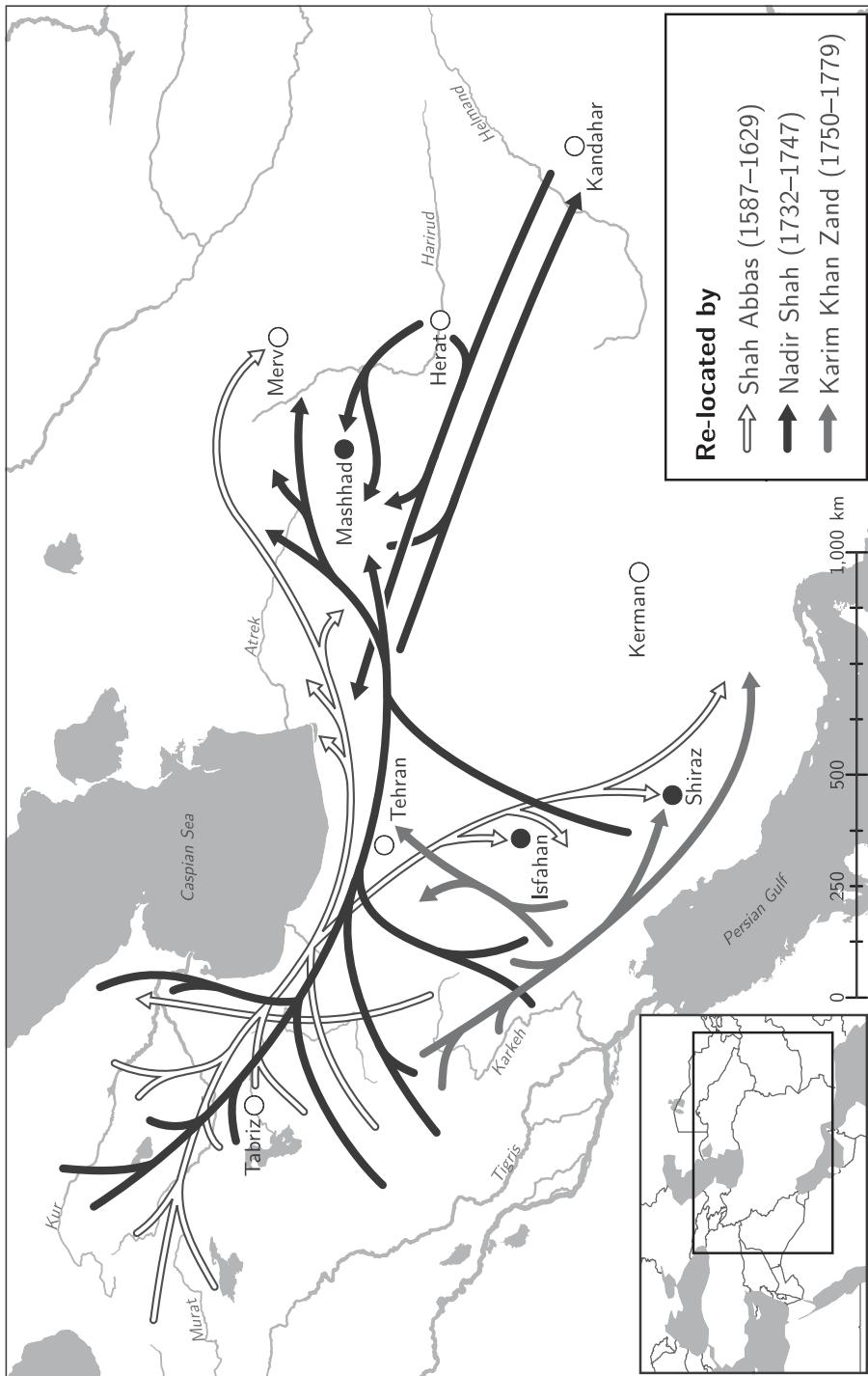


Figure 2: Re-location of nomadic tribes between the 16th and 18th centuries

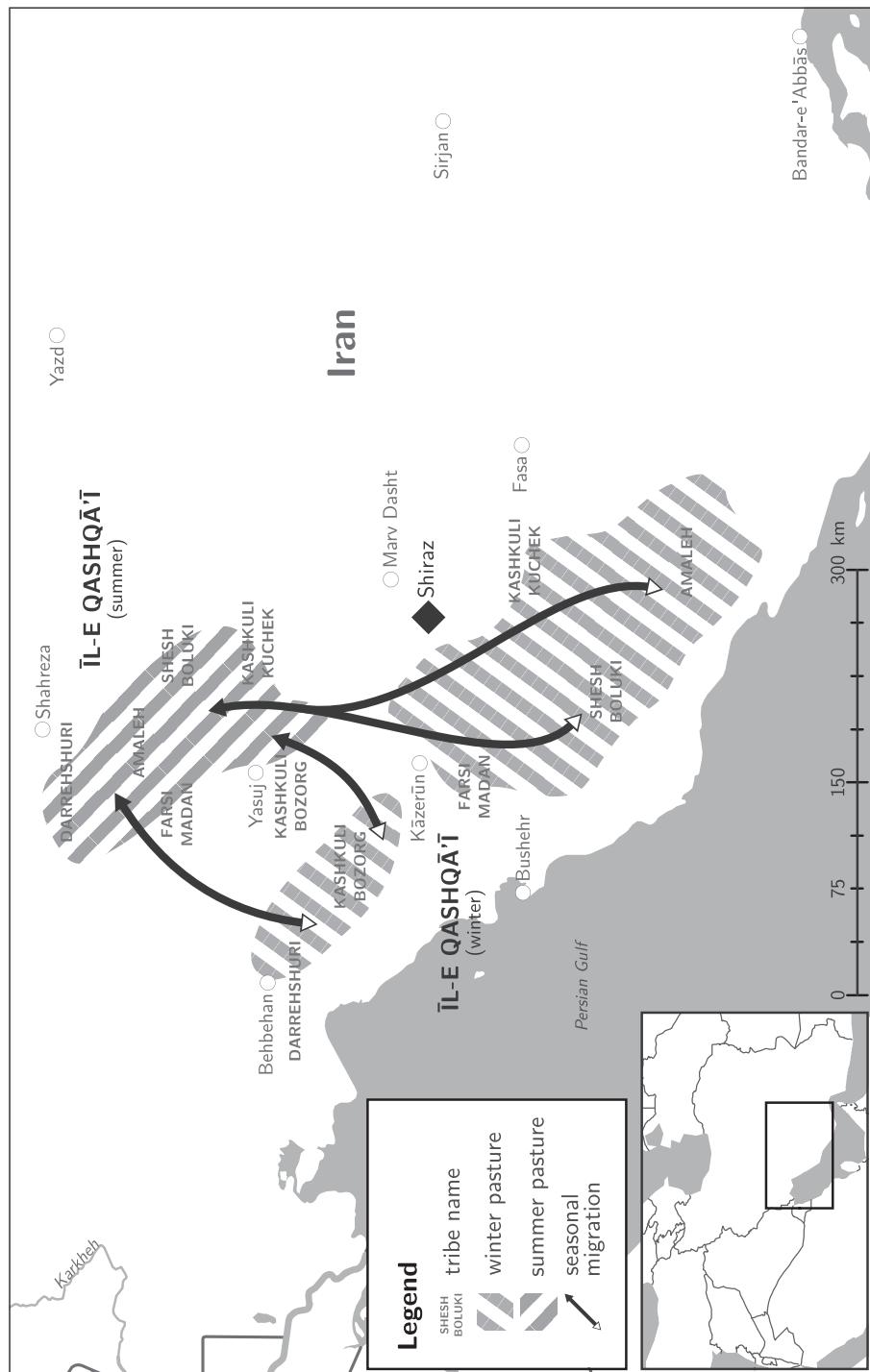


Figure 3: Summer and winter locations of Qashqâ'i tribes

baijan. Outside Azerbaijan, speaking Turkî in public is considered shameful: The usage of a typical low-prestige language implies that the speaker is uneducated, boorish, impolite or obstinate.

Due to wide spread bi- and multi-lingualism, and the reservations some people feel about admitting that they know and use Turkic in some areas of life, it is extremely difficult to estimate the number of actual speakers; and it is even more difficult to figure out the number of Iranians with some competence in Turkic. Moreover, not all the regions where Turkic is spoken, either as the dominant local language or as an inter-ethnic means of communication, have been fully explored.

Doerfer (2006: 93–95) gives some population figures for certain groups of Turcophone Iranians: Azerbaijan (including the exclaves Galûgâh, Lotf-âbâd and Daragaz in Khorâsân): over 12 million; South Oghuz (Qashqâ’î and related dialects): no figure mentioned; Khorâsân (East Oghuz): more than 2 million; Turkmen: 500,000 speakers; Khalaj: (explored in the late 1960s) 28,000. Similarly Boeschoten (1998: 13, probably based on older estimates of Doerfer) assumes that there were 13 million speakers of Turkî in Azerbaijan, 570,000 in the Qashqâ’î region, 400,000 in Khorasan, plus 500,000 Turkmen and 28,000 Khalaj⁴, that is: 14,498,000 speakers of Turkic languages, or about a quarter of the total number of inhabitants of Iran, estimated at 62 million in 1998. The population of Iran today population is roughly 80 million. Given the strong population increase in the rural or tribal areas where the speakers of Turkic languages reside, that would imply that Iran hosts 19 million speakers of Turkî. This figure does not include pockets of Turkic which have still not been explored.⁵ The number of people who use Turkî as a second or third language or as a regional means of communication (and who display different levels of active or passive command of the language) must, however, be much higher.

1.2.2. Bi- and multi-lingualism

Nowadays, most speakers of Turkî are at least bilingual. The education system, improved means of communication and transport, and, last but not least, the omni-present media promote the spreading of Fârsî/Modern Persian as the official language. Thus, mono-lingual speakers of Turkî can only be found among the women of the generation over sixty in regions where Turkî is the dominant lan-

⁴ Assuming that the Khalaj linguistic community has shared the average growth in population figures, Boeschoten (1998) should be corrected to 35,226 out of a total of 78,868,711 according to the recent official population statistics of Iran (July 2011). As the basic estimate of 28,000 was made in the late 1960s/early 1970s, at present the number of Khalaj speakers rather would be over 40,000.

⁵ CIA World Factbook mentions 68 millions inhabitants, of which 24% (16,320,000) are Turcophone.

guage, as, for instance, in Azerbaijan proper, and in rural and tribal areas with predominantly Turkophone populations. And even these mono-lingual women have some passive command of Persian, acquired through contact with their children or grandchildren, or with the media. Men with the same social background are mostly bi- or multilingual, as they have served in the military and are exposed to other languages through work or in their social networks. Normally, they display active command of Persian, plus maybe one or two regional languages of communication.

The social contexts of bi- and multilingualism are very diverse. In Sonqor in the Province of Kordestan, for instance, everyone has some active command of Persian, while Turkî is used as a home language and for trade in the city; the regional language of communication, though receding in recent times, is Kolyâ'i-Kurdish. Armenians from the Central Province or the Esfahan Province speak Armenian, Persian and Turkî, but sometimes additionally also varieties of Kurdish. Kalhor Kurds in the eastern Hamadan Province use Kurdish as a home language, Turkî as the regional language of communication, and speak Persian perfectly well.

1.2.3. Written language

The written language under the Saljuqs was predominantly Persian. Yet, a few literary texts were also composed in *olga-bolga* varieties of Oghuz, or, beginning with the mid-15th century, in the Chaghatai tradition. In the 16th century, poetry in Turkî or “Azeri” starts to develop. This written language of Iran is sometimes called Ajam Turkic.⁶

Although the Constitution of the Islamic Republic grants legal rights for publishing in minority languages⁷, relatively few books in Turkic language are published in Iran. These books mostly contain poetry, folklore texts, collections of fairy tales, religious treaties or research on local history. Publishing concentrates around Tabriz, and to lesser extend, also Qom and Shirâz. In most instances, the Turkic language is represented by an extended form of the Arabic alphabet, which was introduced by the journal *VARLIK* in the 1980s. The linguistic standard is either the Tabriz variety or Turkish of Turkey, which is considered to be “correct Turkic”.

⁶ After the former term *lingua turcica agemica*, which Raphael du Mans coined in the 2nd half of the 17th century, see Johanson (1985: 145).

⁷ Paragraph §15 of the Constitution of the Islamic Republic reads as follows (my translation): “The common official and literary language of the people of Iran is Fârsî. But the use of the local or tribal languages (*zabân-e mahallî va qavmî*) in printed matters and in communication of the respective groups, as well as in teaching of their literature in school, [as a second language] besides Fârsî, is free.”

2. Linguistic features

With the exception of Turkmen and the non-Oghuzic elements in Khalaj, Iran-Turkic varieties are quite similar to one another. They share most of their structural features and their lexicon, and a considerable proportion of elements copied from Modern Persian, which additionally contributes to mutual intelligibility. To exemplify some of the characteristic developments in the Turkic varieties in context, the appendix contains two short sample texts. Text 1 from Sonqor displays a number of individual traces in phonology and morphology, while sample text 2 from Bayâdestân in the Central Province represents main-stream Turkî. Quotations from these texts are marked by a number indicating the text and sentence number. When referring to texts collected in other regions, I will also indicate the place of origin in brackets.

2.1. Phonology

With regard to their phonological behaviour, Iran-Turkic varieties are considerably diverse; consequently, they “sound” different. In most cases, the pronunciation is heavily influenced by the languages in contact of the neighbouring area. Speakers from Iraq or Sonqor have a heavy Kurdish “accent”; the Qashqâ’î pronunciation imitates the local Luri or Shirazi dialects; speakers from the North and the Tabriz region display heavy “fronting”, which is a typical areal feature. Strangely out of place appears the so-called Qipchaq isogloss, a morpho-phonological assimilation well-described for North Azerbaijani, which re-surfacing in the varieties of the Khalaj of Bayâdestân in central Iran (see §5.1.4).

Traditionally phonetic or phonological copying⁸ seems to have been the result of convergence in smaller regional units or of the influence of smaller regional languages of communication, but now it has been widely replaced by replication from the prestige or standard languages that the media propagate.

2.1.1. Consonant inventory

Across the region the basic consonant inventory of the Turkic varieties displays

- five pairs of voiced/unvoiced stops and affricates, namely: the bilabial stops *p/b*, the alveolar stops *t/d*, the velar stops *k/g*, the post-velar/uvular stops *q/χ*, and the affricates *č/c* [ʃ/dʒ].
- four pairs of voiced/unvoiced fricatives: the labio-dental fricatives *f/v*, the alveolar sibilants *s/z*, and the post-alveolar sibilants *š/ž* [ʃ/ʒ], and the velar fricatives *x/y*.
- a voiceless glottal fricative *h*.

⁸ Based on the code-copying model developed by Johanson (1993), contact-induced linguistic processes are perceived as code-copying.

- two nasals: the bilabial *m*, and the alveolar *n*.
- two liquids: the alveolar approximant *r* and the lateral approximant *l*.
- one glide: the palatal approximant *y* [j].

The voiced post-alveolar sibilant *j* [ʒ] is a foreign element that does not appear in genuine Turkic lexical items. Moreover, Turkic words do not begin with *c*- [dʒ], *y*-, *l*-, *m*-, *r*-, and *z*-; the appearance of word-initial *n*- is restricted to interrogatives such as, for instance, (*nimä*) *nä* ‘what’, *niyä* ‘why’, *näcä* ‘how’, and *näčä* ‘how much’. Similar *h*- (< older *q*-) appears in interrogatives *haçan* ‘when’, *harda* ‘where’, and *hangu* ‘which’. Word-initial *m*-, as, for instance, in the personal pronoun 1st singular *män* (< *bän* ‘I’) or the accusative of the demonstrative pronoun (*munu* < *bunu* ‘this’:Acc) has originated through regressive assimilation of *b*- to the following *n*.

The distribution of initial stops is complicated. According to current theories in Turcology, Ancient Turkic had *t*-, *k*- and *b*-, with a tendency of the Oghuz branch to develop *lenes* variants of *t*- > *d*-, and *k*- > *g*-; *p*- rarely occurs word-initially (Doerfer 2006: 97). In present day varieties of Oghuz Turkic the occurrence of *lenes* or *fortes* allophones of *k*- or *t*- varies across different regions, which makes it unpredictable. Most Azeri or Iran-Turkic dialects prefer word-initial *fortes* in Turkic lemmata such as, for instance, *tik*- (corresponding to Turkish *dik*-) ‘to sew’, *tök*- (instead of Turkish *dök*-) ‘to pour’, or *köč*- (for *göç*-) ‘to nomadize’ (but mostly *gäči* for *käči* ‘goat’), while *degil* ‘not’, *gör*- ‘to see’, or *git*- ‘to go’ have *lenes* in both languages. In combination with back vowels, the old *lenes* seem to appear more frequently; the pronunciation of these elements, though, is rather that of a *media lenis* (a devoiced *d* pronounced with strong pressure, resembling the emphatic (pharygealized) Arabic *đ*), see examples such as *d'uz* ‘salt’, or *d'aš* ‘stone’. On the other hand Iran-Turkic or Azeri varieties display word-initial *b*- in a number of lexical items beginning with *p*- in Turkish, see examples such as *barmaq* (in contrast to Turkish *parmak*) ‘finger’, or *p-* in words beginning with *b*- in (Standard) Turkish, as in *pox* (cf. Turkish *box*) ‘shit’. Other Turkic roots vary between *p*- and *b*-, such as, for instance, *piš*- ~ *biš*- ‘to cook’.

Turkic [q] is an allophone of /k/ that appears in +back syllables and occurs in complementary distribution with [k], compare *qul* ‘slave’ and *kül* ‘ashes’. Most Turkic varieties of the target area have preserved word initial *q*-, see, for instance sample text (1/3) [qabər] *qabr* ‘grave’, and (2/6) [qonʃusinən qizenňin] *qonşusunun qızının* ‘his neighbour’s daughter’s’; in some regions *q*- may also be pronounced as postvelar voiced stop [g], which often happens in instances of Sandhi assimilation. Note that Modern Persian pronunciation does not differentiate between the voiced velar fricative /g/ and the unvoiced postvelar stop /q/, as both have merged into a unique phoneme pronounced [ɣ]. Copies of this merger can be found across Iran-Turkic varieties; especially lexical copies from Persian imitate the Persian pronunciation, see, for instance, *qazâ* ‘food’, which also appears with initial [ɣ-] as [ɣæzv]. Even word-initial /q/- in Turkic stems may be pronounced as [ɣ-],

mostly in Sandhi assimilations after voiced consonants or vowels, see the following examples from Bayâdestân in central Iran: [xalam yîzînij] <*halam qızının* ‘my cousin’s’, [bo yejnanasi] *bu qaynanasi* ‘this (was) her mother-in-law’, [de yojmij] *de qoymuš* ‘and has put’, but also [qojær_əmîş] *qoyarmış* ‘one (obviously) used to put’ (Bayâdestân). In word-internal position [q] quite regularly becomes [x], see examples such as *oqu-* [> oxu-] ‘to read’, *aqar* [> axar] ‘it flows’. Especially in copies of Persian lexical items word-internal /q/ is often pronounced as [-y-], see, for instance, the sample text 1 from Sonqor: [motʰæyid], from Persian (< Arabic) *motaqâ’ed* ‘convinced’ (1/3), and [bɒ:yı yalmi] *bâqî qalmış*, a mixed copy of the Persian *bâqî mândan* ‘to remain’, in (1/4). In word-final position of mono-syllabic words /q/ mostly becomes –x, as in *yoq* [> jox] ‘non-existent’, but it also displays the voiced allophone [ɣ], see forms such as [jemay] ‘food’ (Iraq Turkic) or [yatı:y ele-] *qatiq elä-* ‘to prepare dried yogurt’ (Qashqâ’i).

Originally, the members of the pairs of voiced/unvoiced stops *p/b*, *t/d*, and *k/g* did not form separate phonemes in Turkic. Due to the high percentage of bi- and multi-lingual speakers in the region, and the fact that a major proportion of the lexicon has been copied from other languages (such as, for instance, Modern Persian) where */p/, /b/, /t/, /d/, /k/ and /g/ are distinct phonemes, the situation has changed in present day Iran-Turkic. There is a practical need to differentiate between Turkic *bul!* ‘find!’ and the Persian *pul* ‘money’, *gäč* ‘plaster’ and *käč* ‘crooked’ (<*gağ*, *kağ*, both loans from Persian appearing across Iran-Turkic varieties), or Persian *kâl* ‘unripe’ and Turkic *qal!* ‘stay!’*

Table 2: Inventory of consonants (allophonic variants and non-phonemic consonants in loanwords are in square brackets)

	Bi-labial	Labio-dental	Alveolar	Postalveolar	Palatal	Velar	Postvelar/uvular	Pharyngeal	Glottal
Stop	p b		t d			k g	q ɣ	[χ]	[?]
Nasal	m		n		[n]	[ŋ]			
Fricative	[β]	f v	s z	ʃ ʒ		x ɣ	[χ ʁ]	[h]	h
Affricate				f ɖʒ					
Approximant/ glide	[w]				j				
Trill/liquid			r						
Lateral approximant/ liquid			l						

2.1.1.1. Secondary articulation

Aspiration of stops: k, t, p

The unvoiced stops *k*, *t*, *p* are heavily aspirated in syllable initial position and word-finally before a pause, see examples such as [tʰepʰø] ‘hill’, [kʰe:] ‘that’, (both: 1/1) and [bækʰ] ‘beg, prince’ (1/3). The so-called “fronting” (see §4.5) only occurs in the varieties situated north of a line drawn from Qazvin to Zanjan.

Long consonants in intervocalic position

Another characteristic feature shared by the Turkic varieties of the region of Anatolia, northern Azerbaijan, Iraq, and western Iran is the appearance of long consonants. Lengthening mostly relates to stops and alveolar fricatives such as [ss] and [ʃʃ] in intervocalic position, as, for instance, in the following examples from Qashqâ’î: [?æʃʃəg] ‘donkey’, [qʃʃa:yı] ‘downwards’, and [ɪʃʃɪ] ‘hot’. The numerals which regularly display long consonants are [ikkʰe] ‘two’, [jedde] ‘seven’ (also pronounced as [je?de]), [sækkʰɪz] ‘eight’, [doqquz] ‘nine’, and [ottʰuz] ‘thirty’.

2.1.1.2. Copied elements

Pronunciation of ‘ayn [ʃ], bilabial [w] and the pharyngeal fricative [h]

Turkic varieties in contact with Semitic languages or Kurdish dialects —such as, for instance, Iraq Turkic (see the entry on Iraq Turkic: Bulut, this volume, chapter 3.5), Sonqorî, a Turkic variety surrounded by Kolyâ’î, Feylî and Gorani varieties, and, to some extend also the Turkic of the Turcophone Kalhor Kurds in the eastern Hamadan Province—display copied or contact-induced pronunciations of Semitic phonemes, such as the voiced pharyngeal fricative ‘ayn [ʃ], the unvoiced pharyngeal fricative [h], the bilabial [w], and the uvular plosive [q], see examples from Sonqorî, such as [mæħell] ‘place’ (1/2), [ħættʰv] ‘till, up to’, [wa:r] ‘exists’, [wætʰæn] ‘homeland’ (1/4), and also [sæħaq] ‘time’.

Note that these copied phonetic items do not have the status of phonemes; consequently the degree of their realization may vary, and their distribution is not always true to the original.

2.1.2. Vowel inventory

Turkic languages display a basic inventory of 8 vowel phonemes in highly symmetrical distribution: two high/close vowels, namely front *i* [i] and central *ı* [i], and two low/open-mid vowels, namely front *ä* [ɛ] and back *a* [ɑ]. Every unrounded vowel has a rounded counterpart in roughly the same position, forming four pairs: *i* [i]/*ü* [y], *ı* [i]/*u* [u], *ä* [ɛ]/*ö* [ø], and *a* [ɑ]/*o* [ɔ]. Additionally, most varieties of the region have preserved a ninth vowel phoneme — the so-called closed *é* [e].

Table 3: The Turkic vowel inventory

	HIGH/CLOSE				
	i	ü	i	u	
FRONT	(é)				BACK
	ä	ö	a	o	
	LOW/OPEN				

The distinctive primary length of older stages of Turkic has been lost, although one comes across long vowels which coincide with old length (in words such as, for instance, [ɑ:d] ‘name’ or [o:d] ‘fire’, examples from Iraq Turkic).

2.1.2.1. Contact-induced changes

Iranicization

Iranian or Semitic vowel systems, on the other hand, display three distinct vowel qualities, each of which appears in pairs of phonemes distinguished by length (namely *i/i*, *a/â*, *u/û*). Modern Persian — at present the most influential contact language — has lost distinctive length; the 6 vowel phonemes are now *a/â*, *e/e* (<*i/i*), and *o/u* (<*u/û*), while the old quantitative opposition /a/ vs. /â/ has partly developed into a qualitative opposition of front /a/ [a, æ] vs. back /â/ [ɑ, ɔ:, å:]. These additional qualities of /a/ have been copied into Turkic. They appear in copies of Arabic or Persian words, see examples such as: [æl?ɒn] < Persian < Arabic *al’ân* (sample text 1/4) ‘now’, [zæmɒn-i:] < Persian *zamân* ‘time-POSS.3SG’ (1/1), [dær̩s] < Persian < Arabic *dars* ‘lesson’ (2/1), [ɒdæmij] < Persian < Arabic *âdam* ‘man, human being.GEN’ (2/4). In Turkic roots the front half-open [æ] sometimes appears as a very open realization of the vowel phoneme /e/, see examples such as [bækʰ] *bäg* ‘lord, prince’ (1/3), and [gæ:lin] < *gälin* ‘bride’ (2/6).

In Turkic stems, thus, [æ] is in most instances an allophone of /e/, while it represents a realization of short /a/ in Iranian and Semitic roots. With regard to the Turkic phonological principle of palatal harmony, [æ] is of course problematic, as it is situated on the borderline between front and back vowels. Labialized [â], on the other hand, clearly functions as an allophone of Turkic [a].

Delabialization

Another characteristic feature of the so-called Iranization of Turkic vowel systems is delabialization of the front rounded vowels /ö/ and /ü/. In most instances /ö/ is transformed into its unrounded counterpart [e] (as in *öz* > [ez] ‘self’, *böyük* > [bejek] ‘big, grand’, *gör-* > [ger-] ‘to see’), while /ü/ becomes [i], as in *süt* > [sid] ‘milk’, or *yüng* > [jing] ‘wool’. Especially in the vicinity of the glide [j] or palatal [g’] and [k’], [i] < [e] < /ö/ occurs, e. g. *böyük* > [bijek] ‘big’, or *çökäläk* > [çikelig] ‘fresh cheese’. Delabialization is more or less conventionalized in the southern varieties of Iraq Turkic (Bulut 1999), in most Qashqâ’î varieties, in Khalaj in

central Iran (Doerfer 1988), in the Afshar varieties of Qabol and in the dialect of 'Alî Qûrchi south of Arak (see Doerfer and Hesche 1989: 28–29).

In irregular distribution—mostly as spontaneous variants of labialized forms—delabialized pronunciations appear across the whole region, in south-eastern Anatolian dialects with a strong Kurdish adstratum, or with bilingual Turkic-Persian speakers.⁹

2.1.2.2. Intra-Turkic phonotactic processes

Shortening or elision of high vowels in the first syllable

Shortening or even complete elision of a high/close vowel in the first syllable occurs, if this vowel is surrounded by two of the following sounds: stops such as /k/ [k, k^h] and /t/ [t, t^h], which mostly display aspiration as a secondary feature, the unvoiced affricate č [tʃ], and the alveolar fricatives /s/, /š/ [ʃ] or /z/, and the velar fricative [x], e. g. [ʃ'kɪl] < šäkil ‘form’, [ts'xarı] or [tsxar] < čixar(i) ‘he/she comes out’ (Iraq), [tʃʃp:x] < ušax ‘child’, [tʰkɛ:rde] < tikärdi ‘was sewing’ (Sonqor), [tʰtalla'] < tutarlar ‘they hold’ (Bayâdestân), and, with a Persian root, [b'lstærsi] < bištar.POSS.3SG ‘most of it’ (Esfahan).

This intra-Turkic phonological process, which appears across the Turkic varieties from northern Azerbaijan, Iraq, and western Iran obviously is related to the internal structure of the syllable: unvoiced/fortis consonants, especially when combined with aspiration, are longer than voiced, that is, in most instances, lenis consonants. To avoid overlong syllables, the fortis seem to acquire some of the length of the vowel, which consequently becomes shorter.

Metathesis and rhotacism

Throughout the region, different forms of metathesis, lambdacism or rhotacism appear, as in examples such as [torpax] < topraq ‘ground, soil’, [jarpax] < yapraq ‘leaf’, [tʃilpax] < čiplaq ‘naked’, [irəli] < iläri ‘forward’, [kørþy] < köprü ‘bridge’, etc. Consequently, stems and morphemes containing /l/ or /r/ are most susceptible to various kinds of metathesis or assimilations (see below, morpho-phonology). The consonant group /-lq-/ is pronounced as [x], both in Turkic or in loaned items, e. g. /-lq-/ > [x]: in [qax] < qalq- ‘get up’, or [xax] < xalq ‘people’.

2.1.2.3. Contact-induced phonotactic developments

Instability of initial y-

A typical Azeri isogloss is the loss of word-initial y- before high/close front vowels such as, for instance, *i*, *ı*, and *ü*. Across Türkî varieties of the region there is a ten-

⁹ Doerfer makes similar observations regarding Paradomba (Esfahan province), see Doerfer and Hesche (1989: 28–29).

dency to delete old *y*- in front of *i*, as in *igit* ‘hero’ (corresponding to Turkish *yığıt*) and *itik* (< *yitik*) ‘lost’, or to replace the sound group *yi-* by *i-*, as in *ilan* ‘snake’, *il* ‘year’ (< Turkish *yılan*, *yıl*). The same varieties may display variation, e. g. in forms like *yüng* ~ *üng* ‘wool’ used interchangeably in Qashqâ’î. More unusual is the development of *yo-* > *o-*, as in *yoxari* ‘upwards’, which has the form *oxari* in most Qashqâ’î varieties. In other instances, initial *y*- is preserved, see examples such as *yiy-* ‘collect-’, *yrylamax* ‘to weep’ (Sonqorî), or *yit-* ‘to loose’. Note that word-initial sound groups such as *yi-*, *yi-*, or *yü-* are alien to the contact languages, such as Modern Persian, where in genuine Iranian words only *ya-* or *yā-* occur.

Initial h-

Across the Turkic varieties of the region initial *h*- is problematic. In some lexical items, such as *his* ‘soot’ (cf. Turkish: *is*), *hör-* ‘knit’ (Turkish: *ör*), or *hürk-* ‘shy’ (Turkish: *ürk-*), it coincides with older Turkic *h*-, which has been lost in other Turkic languages. There are also numerous examples of conventionalized prothetic *h*- in connection with foreign words, such as *helbet* (< Arabic *albatte*) ‘naturally’, or *häqiq* (< Persian < Arabic *aqīq*) ‘agate-stone’. In other instances spontaneous *h*-prostheses occurs, e. g. *härwär-äkä* ‘the woman’ (< Arabic ‘awrat-SPEC) from Sonqor, or *helämijälär* ‘they don’t do’ (< *elämiyirlär*) from Bayâdestân. Prothetic *h*- may have originated under the influence of contact languages such as Arabic or Kurdish, which do not display word-initial vowels. Turkic varieties in different language-contact situations completely avoid *h*-, e. g. *äm* ‘too, also’ (< Iranian *ham*), or *är* (< Iranian *här*) in the Urum variety of Georgia.

2.1.2.4. Replicated areal features

Fronting

One of the most characteristic distinctive features in the pronunciations of the Turkic varieties of the region is the so-called “fronting” (Stilo 1994), a phonological process denoting palatalization of the velar stops /k/ and /g/ to [k̯] and [g̯], or, in a more advanced form, affrication to [dʒ] or [ʃ], in which case the original affricates [dʒ] and [ʃ] are fronted to [dz] and [ts/tç], see Bulut, this volume, chapter 3.5, §2.1.1. This areal phenomenon is found in the pronunciation of genetically unrelated languages across Iraq, Anatolia, Azerbaijan and Iran. It is most prominent in the Turkic varieties of the Azerbaijani exclave Nakhichevan, in Qarabagh, in some regions of Georgia, around Tabriz and Urmia in Iran, and in northern Iraq (Bulut, this volume, chapter 3.5).

Zagros-d: /-d-/ > [-r-], [-j-]

In Kurdish and Goranî dialects of the Zagros area, intervocalic -d- is unstable (see Mann 1930: 96). Sonqorî, a Turkic dialect of this region, has copied the instability of -d- as an areal phonological feature. The so-called “Zagros-d” becomes [-r-]

with back and front vowels, and sometimes [-j-] in combination with front vowels. This happens in connection with copied lexical items such as *âdam* [aræm] (< Persian < Arabic), ‘man’, and *adaqlı* [araxli] ‘fiance’, as well as with genuine Turkic roots, such as *odun* [urın] ‘firewood’, or with the copula past 3SG in *idi* [iri] ‘was’, and present 3SG in {+ri} < {+dir} ‘is’, see sample text (1/2), or verb forms such as *gedärdi* [gijärdı] ‘he/she went along’.

To a lesser extent instability of intervocalic /-d-/ also appears in Qashqâ’î in the southern Zagros area. In connection with the verb stem *ged-*, for instance, intervocalic /-d-/ may either be pronounced as [d], or it may be replaced by the glottal stop [?] or the palatal glide [j], as the following forms based on the verb root *ged-* ‘to go’ demonstrate: [ge?illær], [gejllær], [gejillär], all three forms being pronunciations of *gedirler* ‘they go’; see also Bulut (2016: 250–251).

South-western /b/ > [-v-] ~ [-w-]

The most characteristic areal feature in the phonology of Qashqâ’î varieties are secondary diphthongs or long monophthongs, which appear with foreign and genuine Turkic words. In intervocalic position, especially in combination with open vowels such as /a/ or /o/, /b/ becomes > [v], or even [w], while /v/ develops into the voiced bilabial [w]. In a second step, the resultant sound sequences, such as [awa], [awä], or [uwa] etc. may be contracted to diphthongs such as [oa], [ou], or develop into a long monophthong, such as, for instance, [o:]; e. g. /-b-/ > [-v-] in [urdiveheʃd] < ‘*Ordibehešt*’ (Iranian month); /-b-/ > [-v-] > [-w-] in [owa] < *oba* ‘camp group’, or [owolarina] ‘camp group’.POSS.3PL.DAT; /-aba-/ > [-oa-] in [boam] < *babam* ‘my father’; /-ava-/ > [-uwa-], [-o?a-], [-oa-], [-oha-], or [-o:-] in pronunciations such as [duwvr] < *davar* (< *tabar) collective ‘sheep and goats’, [do?ar], [doha:r], or [do:rları] < *davar*.POSS.3PL. A special case is [tʰoussən] < *tâbestân* (< *tâwessân < *tâwestân) ‘summer’, which is conventionalized as an item of the Turkic lexicon. Speakers of Qashqâ’î usually do not connect this to its Persian source word *tâbestân*, which they pronounce [tʰøbesdan].

2.2. Morpho-phonology

2.2.1. Assimilations on morpheme boundaries: an Intra-Turkic feature

The following types of consonant assimilations are found across the Turkic varieties of the region.

2.2.1.1. Assimilation of consonant groups containing the liquids /r/ and /l/

/-nl-/ > [-nn-]: Suffix-initial /l-/ is subject to various types of progressive assimilation on morpheme boundaries. The plural suffix {+lar}, for instance, may assimilate to stem final /-n/ as in [dzejrannær] < *ceyranlar* ‘gazelles’ (Iraq), [mehmən-

ner] < *mehmanlar* ‘guests’ (Sonqor), [sazımənner] < *sazemanlar* ‘organizations’ (Esfahan), and pronominal forms, such as [unnar] < **onlar* ‘they’, or [bunnar] and [mənnar] < **bunlar* ‘these’ (Esfahan).

/-šl-/ > [-ʃʃ-]: Suffix-initial /l/ may also assimilate to preceding /-š/, as in [diʃʃere] < *dişləri* ‘her teeth’ (Sonqor). **/-sl-/ > [-ss-]:** occurs stem internally, too, see forms such as [æssæn] < *aslan* ‘by no means’ (Bayâdestân, Esfahan).

/-nl- or /-rl-/ > [-ll-]: The reverse process, namely regressive assimilation to /l/ after stems ending in /-n/ or /-r/ is also frequent; e. g. /-nl-/ > [-ll-] in [ortsillær] < *örtsünler* ‘they should cover’ (Iraq); /-rl-/ > [-ll-] in [verillər] < *verirlär* ‘they give’ (Khalaj of Bayâdestân), or in [vallərə] < *varlari* ‘they own/possess’ (Esfahan).

/-ld-/ > [-ll-]: Combinations of /-l-/ and /-d-/ may display progressive assimilation at morpheme boundaries; e. g. [qallire] < *qaldırır* ‘he picks up’ (Iraq), or [ollu] < *oldu* ‘he/she became’ (Esfahan).

2.2.1.2. Assimilation of consonant groups containing the nasal /n/

/-nd-/ > [-nn-]: This type of assimilation occurs stem-internally in Turkic and Persian roots; e. g. [kìnne] < *künde* ‘dung cakes’ (Bayâdestân), [tænnor] < *tandur* ‘oven’ (Khalaj of Bayâdestân), and [zənnigisənne] < *zändegisindü* ‘in her/his life’ (Sonqor) — as well as at a morpheme boundary, as in, for instance, [sallannə] < *sallandi* ‘was pending’ (Khalaj of Bayâdestân). **/-nd-/ > [-nn-]:** is frequent in combinations with the case suffixes of the locative and ablative. There is, however, a great deal of variation: In Iraq, the initial /d-/ of the locative is often preserved, e. g. forms such as [Pattanda] ‘while throwing’, while the ablative mostly exhibits assimilation [alt'_jaſınnan sora] < *altı yaſından sora* ‘after the age of six’. Often, assimilated forms of the locative appear alongside non-assimilated ones, as in [janında] < *yanında* ‘by his side’, and [açınna] < *ačanda* ‘while opening’, both pronounced by the same speaker.

2.2.1.3. Assimilation of alveolar stops to sibilants

/-zd-/ > [-zz-], /-st-/ > [-ss-]: Similarly, the alveolar stop /d/ and its unvoiced counterpart /t/ assimilate to the preceding sibilants /s/ and /z/. This type of assimilation, which is also frequent in spoken Persian, may occur stem-internally with Iranian roots; e. g. [døesse] < Persian *daste* ‘bunch, set’ (Sonqor), [bis] < Persian *bist* ‘twenty’ (Esfahan). With Turkic lexical items, a similar assimilation may also appear at a morpheme boundary; e. g. [bɒʃərməzzim] < *baʃarmazdım* ‘I did not succeed’ (Sonqor).

2.2.1.4. The Qipchaq isogloss (-> Regional feature)

A special type of assimilation of the plural suffix $\{+lAr\} > \{+DAr\}$ appears with speakers of the so-called Khalaj tribe in Bayâdestân¹⁰ region. The initial *l*- of the plural suffix becomes $> [d-]$ after stem-final *-z*, as in [øzdaer] < *özler* ‘they themselves’, and $> [t-]$ after stem-final *-š* and *-t*, as in [dijermiſter] < *diyärmiſlär* ‘they used to say’, or [zattʰær] < *zatlär* > ‘things’. This kind of */l-/ > [D-]* assimilation is generally associated with the Qipchaq language group; yet, it also occurs frequently in North Azerbaijani dialects (see Širaliev 1983: 104, 130 etc.). For Khalaj in central Iran Doerfer (1988: 60) quotes a single example of this type, namely: *tuzla* > *tuzda* ‘with salt’.

2.2.2. Realisation of velar /ŋ/ as [ŋ, y, w] as an intra-Turkic isophone

The development of the old velar /ŋ/ in stem internal position is not predictable. Across the region /ŋ/ has the realization [ŋ] in a number of lexical items, such as, *qaranlux* ‘darkness; cf. Older Turkic *qarangu* ‘dark’, but Turkish and North Azerbaijan *qaranlıq*, that is /ŋ/ > [n]; *yäŋi* (< *yayı*) ‘new’, *yüŋg/iüŋ* etc. (< *yuy*) ‘wool’. In other instances /ŋ/ has become [y], as in *yayluz/yalyuz* (< *yal(1)njuz*) ‘alone’. Iraq Turkic also displays /ŋ/ > [w], as in *gewil* ‘heart’ (< *göyük*), or /ŋ/ > [g], as in < *ög* ‘front’ (< *öŋ*).

Different realizations of old /ŋ/ > [ŋ, y, w] in various different morpho-phonological contexts constitute a prominent feature that distinguishes the Turkic varieties of the region, e. g. in the genitive suffix $\{+nIŋ\}$, or /ŋ/ > [ŋ, y, w], the possessive suffixes of the second person singular and plural $\{+(I)ŋ/\{+(I)ŋIz\}\}$, and, consequently also in copula/the personal suffixes of the possessive and (later also) the pronominal conjugation form. The shift /ŋ/ > [w] appears more frequently in the northern part of the region, in North Azerbaijani dialects and in South Azerbaijan (i. e. the Iranian province of Azerbaijan with the capital Tabriz), and in north-eastern varieties of Iraq Turkic. The southernmost representative of the w-group seems to be the small Turkic enclave of Sonqor, north of Kermanshah in western Iran. The development /ŋ/ > [y] is more characteristic of the central and southern varieties, such as the dialects of Bayâdestân in the Hamadan and Saveh provinces and the varieties around Esfahan, and also the Qashqâ’î varieties in the province of Fars in south-west Iran¹¹. The Saffi Khânî variety of Qashqâ’î, however, has preserved /ŋ/ as [ŋ], see Bulut (2016).

¹⁰ One of the Oghuz-speaking Turkic tribes in the valley of Bayâdestân between Sâveh and Hamadan in central Iran also call themselves Khalaj. Their genealogical relation to the speakers of the archaic Khalaj language is not clear.

¹¹ That {y < * ŋ} is also found in Urfa in eastern Anatolia may be due to the presence of the Bayat tribe in this region.

2.3. Morphology

Turkic morphology is agglutinative and right-branching; derivation and inflection rely on suffixes. Postpositions or case markers combined with postpositions denote location.

2.3.1. Nominal morphology

2.3.1.1. Plural

The plural morpheme {-lAr} is added to nominal stems before possessive and case suffixes. It is subject to certain assimilations, see §2.2.1.1.

2.3.1.2. Possessive suffixes

The possessive suffixes are: 1SG {+(I)m}/1PL {+(I)mIz}, 2SG {+(I)ŋ}/2PL {+(I)ŋIz}, and 3SG {+(s)I}/3PL {+lArI}, with considerable phonological variation, assimilations and contractions in the forms of the 2nd persons.¹² Basically, present day varieties either display combinations of high unrounded vowel /I/ + y > [+Iy] or high rounded vowel + v/w > [+Uw], compare forms such as *ad-uw/ad-iy* ‘your name’, *äl-üw/äliy* ‘your hand’. For the geographical distribution of this characteristic morpho-phonological change see §2.2.2 above. In most Qashqâ’î varieties in the province of Fars in south-western Iran the possessive suffixes of the 2nd persons are {+ɪŋ} ~ [+ɪj] in the singular and {+ɪŋɪz} or [+i:z] in the plural, while Saffî Khânî displays old [ŋ] in all instances. An intermediary stage with palatalized [ɲ] also appears.¹³ Presumably, the phonological development is {+ɪŋɪz} > {+ɪŋɪz} or {+ɪjɪz}, with palatalization of /ŋ/ due to the surrounding close front/central vowels [i] and [ɪ], followed by a contraction to > [+i:z].

¹² The basic forms of the Oghuz possessive suffixes are 1SG {+(U)m}/1PL {+(U)mUz} ~ {+(I)muz}, 2SG {+(U)ŋ}/2PL {+(U)ŋUz} ~ {+(I)ŋuz}, 3SG {+(s)I}/3PL {+lArI}. Doerfer (2006: 102) supposes that the forms of the second persons go back to a mixture of {+(U)ŋ}/{+(U)ŋUz}, with old /ŋ/ > [n], [ŋ], or [ɲ], and {+(U)G}/{+(U)GUz} (only in inscriptions, see Gabain 1941: 97), where old /G/ became [v], [y], [g], [ɣ], or [ŋ]. Note that the genitive < *{+nUŋ} shows the same phonological varieties without being traced back to “old G” < *{+nUG}. Writing /ŋ/ as <G> may have been a graphic convention; see also the representation of /ŋ/ as *sağır kaf* in Arabic script.

¹³ According to Tuna’s notation (1984), palatalized [ɲ] seems to be the most frequent realization of old *ŋ/ in Abivardi.

2.3.1.3. Case morphology

Across the region case morphology displays the typical southwestern Oghuz characteristics, with the following basic forms: Genitive {+(n)Iŋ}, dative {+(y)A}, non-specific accusative {+Ø}/specific accusative {+(n)I}, locative {+DA}, and ablative {+DAn}.

The genitive {+(n)Iŋ} exhibits a number of morpho-phonological variants: {+(n)iŋ} in varieties which preserve the velar nasal, see Qashqâ'î [xvñiŋ qaqası] *xân-iŋ qaqasi* 'the brother of the Khan', or [ezəniŋ] öz-ü-nün 'his/her own'; in the central dialects which display /ŋ/ > [y] it has the forms {+niy} ~ {+(n)Iŋ} (in sandhi before k-, g- and q-) ~ {+(n)In}, e. g. [za:nij] *zan-iy* 'the woman's', [kæslærinj ki] *kæs-lær-iŋ ki* 'of the ones who' (Bayâdestân), and both [kʰæsən] *käs-in* 'someone's' and [pdæmij] *âdam-in* 'the guy's' in (2/4); or {+(n)In} as in Turkish of Turkey, in the northern Iran-Turkic varieties and in Iraq Turkic. Especially in connection with personal names, genealogical/kinship terms or indefinite pronouns, the possessor in a genitive-possessive construction often does not display genitive marking; e. g. *härkäs šärbäti* (everyone.Ø lemonade.POSS.3SG) instead of *härkäs-in šärbäti* (everyone.GEN lemonade.POSS.3SG) 'everyone's lemonade' in example (24) below; *hala-m qız-i* 'my cousin' (aunt-POSS.1SG.Ø daughter-POSS.3SG), instead of *hala-m-in qız-i* (aunt-POSS.1SG.-GEN daughter-POSS.3SG), from Bayâdestân, or *Ahmad ārwâr-i* [Ahmad.Ø wife-POSS.3SG] 'Ahmad's wife', from Sonqorî. Note that heavily Iranicized varieties, such as, for instance, Sonqorî, have a strong tendency to replace the Turkic genitive-possessive construction by the Persian *ezâfe*-constructions, see sample text 1 in the appendix.

The possessive dative is {+(n)A}, that is: [+A] after consonant, and [+nA] with pronominal /n/ in combination with the possessive of the 3rd person, as in *baš-i-na* (head-POSS.3SG-DAT) 'on his/her head'.

The marked accusative in {+(n)I} has the possessive forms {+n} ~ {+nI} (like Turkish of Turkey).

With the exception of Iraq Turkic (where the locative suffix remains {+dA}), the locative assimilates to preceding nasals: {+DA} > [+nA]; when in combination with the possessive of the 3rd person pronominal suffix, *n* appears, as in *göz-däri-n-nä* (<*göz-läri-n-dä* eye-POSS.3PL-n-LOC) 'in their eyes'. In connection with the ablative, the same assimilation {+DAn} > [+nAn] is found across the region. In Sonqorî, the ablative morpheme has lost its final /-n/, see *bannâ-sin-nä* (building-POSS.3SG-ABL) 'of their buildings', see sample text no. (1/1).

Across the region the instrumental case is {+(I)nAn}, as in sample text (2/3): [dʒinner_ənan] *cin-lär-inan* 'with the jinns', while, as a rule, the form {+(y)dan} is characteristic of Iraq.¹⁴ The equative case is based on {+cA}; pronominal [n] does not appear; cf. forms such as [dvlidža] *dal-i-ca* (< back-POSS.3SG-EQU) 'behind

¹⁴ Sporadically one also comes across an instrumental case in {+(y)dan} in Iran, as, for

him/his back' (Bayâdestân). An exception is the equative of the demonstrative pronouns, e. g. forms like [ondza] 'that much'. The morpheme of the terminal case ('up to, till') is {+(y)AcA}, as in [jeredʒe] *yer-äcä*; 'up to the place' (Bayâdestân).

2.3.1.4. Comparison

Comparison of adjectives is expressed analytically by the adverbs *daha* 'more' and *lap* (restricted to Azerbaijan) or *än* 'most' preceding the adjective. The Turkic suffix {+rAK} has survived in most regions of Azerbaijan proper and in Sonqorî. Parallel, many varieties use the Iranian comparative suffix {+tär} and the superlative suffix {+tärin}, e. g. Sonqorî *čox-raq* and *čox-tär* 'more'. In the South, Iranian comparative suffixes have nearly completely replaced their Turkic counterparts.

2.3.2. Pronouns

The personal pronouns are 1SG *män*, 2SG *sän*, 3SG *o*, 1PL *biz*, 2PL *siz*, 3PL *olar* (~ *ular* etc.); the reflexive or emphatic pronoun is based on *öz* 'self'. In contrast to Turkish, there are only two demonstrative pronouns, namely *bu*, Pl. *bular* 'this' (with an emphatic form *ha bu* 'this one here'), and *o*, Pl. *olar* 'that'. While the indefinite pronoun is mostly *bir* 'one', Sonqorî and Ben use a special form *bice* (see 1/1).

Across Turkic varieties of Iraq, and northwestern, central and southern Iran (including Khalaj, see Doerfer (1988: 107) a defective pronominal paradigm based on the stem *bilä-* appears.¹⁵ This always combines with possessive suffixes and case markers, as, for instance, in *bilä-mız-ä* (*bilä* + POSS.1PL + dative) 'for/to us', or postpositions, as in *bilä-si* 'čün' (*bilä* + POSS.3SG + POSP) 'for him/her'. It neither appears as an unmarked case or nominative, nor in the function of adjectival *öz* 'own'. *bilä* is restricted to object positions of dative, accusative, ablative and instrumental, and combinations with the postposition *ičin* 'for'; it does not combine with genitive suffixes to form possessive pronouns (see also Bulut 2003).

instance, in the variety of the city of Ben, in the North-east of Chaharmahal and Bakhtiari Province.

¹⁵ The lexical meaning of this stem is unclear; the best guess is Soper (1996), who maintains that *bilä* goes back to the adverb *belä* 'so', which indeed may combine with possessive suffixes.

Table 4: Pronominal paradigms

	Personal	Personal:DAT	Possessive	Reflexive	Remoteness/ modesty:DAT
1 SG	<i>män</i>	<i>mänä</i>	<i>mänim</i>	<i>özüm/özim</i>	<i>bilämä</i> ('to me')
2 SG	<i>sän</i>	<i>sänä</i>	<i>säniy/sänuw</i>	<i>öziy/özüw</i>	<i>biläyä</i> ('to you')
3 SG	<i>o</i>	<i>ona</i>	<i>onij/onuy</i>	<i>özi/özü</i>	<i>biläsinä</i> ('to him/ her/it')
1 PL	<i>biz</i>	<i>bizä</i>	<i>bizim</i>	<i>özümüz</i>	<i>bilämizä</i> ('to us')
2 PL	<i>siz</i>	<i>sizä</i>	<i>siziy/sizuw</i>	<i>özüyiz/özüwüz</i>	<i>bilayızä/biläyüzä</i> (‘to you’)
3 PL	<i>olar</i>	<i>olara</i>	<i>olary</i>	<i>özläri/özdäri</i>	<i>bilälärinä</i> (‘to them’)

2.3.3. Verb morphology

Turkic verbs express diathesis, time/aspect/mood, possibility and agent by suffixes attached to the stem in this order.

2.3.3.1. *Copula*

Across the area, the heterogeneous forms of copula and possessive suffixes function as a distinctive feature of the Turkic varieties, see Table 5 below.¹⁶ While the copula suffixes of the first and second persons derive from the personal pronouns, the third persons are either unmarked {-Ø}, or take the form of a suffix in {-Dir/-Dur etc.} that goes back to the verb *dur-* ‘to stand’. Frequent usage of the copula of the 3rd person may point to the influence of Iranian languages, where an overt marking is obligatory. The plural suffix {-lAr} quite regularly displays regressive assimilation of word-final [-r] to the [l-]; forms like *olurlar* ‘they become’, are pronounced [olullar] in most Iranian and Iraq-Turkic dialects.

¹⁶ In the tables below, <X> indicates +/- rounded high vowels such as {X = u, ü, i, ī}; <I> represents unrounded high vowels {I = i, ī}, <U> rounded high vowels {U = ü, u}, and <A> open unrounded vowels {A = a, ä}.

Table 5: Copula forms

Old Anatolian Turkic	Standard Turkish	East Anatolia/ Diyarbakir	Standard Azeri	Iraq	Sonqor	Bayâdestân (Safi Khânî)	Qashqâ'i
1 SG -vAn, -Am	-Am -(y)Xm	-(y)Am	-(y)Am	-(y)Am	-(y)Am	-(y)Am	-(y)Am
2 SG -sAn	-sXn	-sAn	-sAn	-sAn ~say ¹⁷	-sA ~saw	-(y)Ay	-(y)Aŋ
3 SG -Durur > -DUr	-Ø/-DXr	-Dir	-DXr	-DX	-Ø/-dI/-rI	-Ø/-dX(r)	-Ø/-dI(r)
1 PL -(v)Uz	-(y)Xz	-(y)iX	-(y)XK	-(y)Xx/ (y)Xy	-(y)AK	-(y)XK	-(y)AK
2 PL -siz	-sXnXz	-siz	-sXnXz	-sXz	-sIz	-(y)X:z	-(y)AŋIz
3 PL -DUr (-lar)	-(DXr)lAr	-(Dir)lär	-(DXr)lAr	-(DX)lAr	-(DI)lAr	-DX(lA(r))	-DI(lA(r))

Especially in the varieties of central and south-western Iran, the personal suffixes of the possessive and pronominal type have become more similar to each other in the 1st and 2nd persons and are only differentiated by a high/close (in possessive suffixes) or an open (in pronominal suffixes) vowel. This differentiation has already collapsed in some of the Central dialects, where — due to internal assimilation to /y/—the 2nd persons may display identical pronunciations, see Table 6 below.

Table 6: Personal suffixes of the possessive and pronominal type/South and Central

	1 st singular	2 nd singular	1 st plural	2 nd plural
possessive type	{-Im}, {-Xm}	{-Inj} > {-iy}	{-Ik}, {-Xk}	{-InjIz} > {-I:z}
pronominal type/ copula	{-Am}	{-Aŋ} > {-Ay} > {-ey} > {-iy}	{-Xk}, {-Ak}	{-AŋIz}, {-I:z}

2.3.3.2. Tense-mood-aspect

Focal present and aorist

Across the region, most present paradigms reflect the Azeri type: The morpheme of the focal present, i. e. progressive present, is {-Xr}, which differs from the aorist in {-Ar} only with regard to its high/close vowel. Stem-final vowels are

¹⁷ In Kifri, Tuzkhurmati etc.

normally deleted; thus, for instance, *işlä-* ‘to work’ forms a focal present *işl-ir(i)* ‘he is working’, and an aorist *işl-är* ‘he (usually) works’.¹⁸

Deviations from this model are frequent. In varieties that display a morphologically different focal present, high vowels such as {X = u, ü, ı, i} may also appear in the aorist, their distribution being similar to that of Turkish of Turkey.

Sonqorî uses a focal present in {-({X}(y)ou(r))}¹⁹, e. g. *issyouräm* (also contracted: *isiyoum*) ‘I want’, *bilowsä:* ‘you know’, *iliyou* ‘he/she is doing’, *issyourax* ‘we want’, *bilousez* ‘you (PL) know’, and *issiyouse:z* ‘you (PL) want’, and *issiyoular* ‘they want’. Consequently, high/close and open vowels appear in the aorist, as in the following examples from sample text 1: *danifille:* speak.AOR.3PL (1/1), *di jelle:* say.AOR.3PL (1/2) *ilelle:* make.AOR.3PL (1/5). In other varieties, relics of a present construction in {*-Iyori} occur alongside the more frequent forms based on {-Ir}, e. g. *giriyöm* (<*gör-iyür-äm*) ‘I see’, and *bilmiyüräm* ‘I don’t know’, both from the Bayat variant of Fars, or *qalyür* ‘he stays’ from the Khalaj of Bayâdestân.

Due to a tendency across the region to shorten or delete word-internal syllables, the distinctive feature of the present or aorist morpheme ({-Xr} vs. {-Ar}) is often blurred. Moreover, in the 1st person singular and plural the vowel of the aorist is regularly omitted after stems ending in /-r/ or /-l/, e. g. *vur-r-am* (<*vur-ar-ram*) ‘I beat’ or *allam* (<*al-ar-am*) ‘I take’; see also Bulut, this volume, chapter 3.5, §2.1.2.1.

Personal suffixes of the present

The personal suffixes of the present represent the so-called pronominal type of conjugation. Thus, the agent markers of the first and second persons of the present are identical with the copula forms, with the same phonological variation, especially in the second persons. Across the area, the suffix of 1st person singular is {-Am}, as in *bilir-äm* ‘I know’; the 1st person plural is based on {-IK} in most northern and central varieties, as, for instance, in *säläš-ir-ig* ‘we speak’ (Iraq), or *ališirix* (Bayâdestân) ‘we learn’; in the South {-AK} also appears, as, for example, in *iss-your-ax* ‘we want’ (Sonqor), or *biz gäl-ir-äk* ‘we come’ (Bayat of Fars).

In the northern varieties, the 2nd persons have the forms {-sAn} (or rarely also: {-sAy}) in the singular and {-sXŋXz}, {-sX:z}²⁰, or {-sXz} in the plural, e. g. *bilirsän* ‘you know’, or *alisız* ‘you.PL take’, with elision of the /-r/ of the present marker in Iraq Turkic.

¹⁸ An exception to this rule are monosyllabic verb stems ending in vowel, such as *ye-* ‘to eat’ and *de-* ‘to say’, which may, among other possible forms (such as *deri*, *deyiri*, *diyiri* etc.), have a present tense *deyir(i)* ‘he says’, and an aorist *diyär* ‘he (usually) says’.

¹⁹ Like the Turkish focal present, this tense marker is based on *{-({X}yori)}, consisting of a gerundial form of the lexical verb plus an auxiliary < *yori- ‘to go’.

²⁰ Contraction of {-sŋiz/-sŋiz} or {-seniz/-saniz}.

In the central and southern varieties of the region, the personal suffixes of the 2nd person singular and plural go back to an older intermediary form {-Aŋ} or {-AŋIz}, which the Saffî Khânî-variety of Fars has obviously preserved, e. g. *sän gäl-ir-äŋ* ‘you.SG come’ or *siz gäl-ir-äŋiz* ‘you.PL come’. Other varieties from the Îl-e Qashqâ’î or from central Iran exhibit different stages of the phonological development of {-Aŋ} > [-Ej] > [-ej] > [-ij]. The Begdilli and Bayat variants of Fars, and the central dialects of the Hamadan Province and the region around Esfahan have *gäl-ir-äy*, or rarely also *gäl-ir-iy*, in the singular, while the corresponding forms of the 2nd plural are *gäl-ir-äyz* or *gäl-ir-eyz* (with an additional variant *gäl-ir-iyz* in the Bayat variety of Fars and some central dialects).

Across the region, the 3rd person singular appears without an agent marker, as, for instance, in *gör-ör* ‘he sees’ (2/6); yet, it also may exhibit an additional element {-I} = [i, e], as in *öltre* ‘he/she dies’ (Iraq), or *isture* ‘he/she wants’ (Bayâdestân), or *der-e* ‘he says’ in sample text 2/4. The final /-r/ of third person plural often assimilates to the initial /l-/ of the plural suffix, as, for instance, in *gäl-il-lär* (come.PRS.3PL) ‘they come’ (2/4), *de'illär* (say.PRS.3PL) ‘they say’ (Esfahan), *de-'äl-lär* (say.AOR.3PL) ‘they (usually) say’ (2/5).

The functional distribution of the aorist (habituality, iterativity, ability, intention) and present (focality) is not clear-cut. In Standard Azerbaijani and in the northern varieties, it mostly corresponds to that of Turkish of Turkey. As explained above (see §2.3.3.2 on *Focal present and aorist*), the central dialects and Iraq Turkic sometimes exhibit a formal merger of the positive forms, which makes it impossible to differentiate between aorist and present. The functional distribution of different paradigms which may exist in parallel in the same variety — such as the present in {-Ir} and relics of a present in < {*-Iyori}> in the Bayat variety of Fars — is also unclear. In some Qashqâ’î-varieties the present in {-Ir} has lost focality, in that its function has been extended to include not only the focal present but also iterativity or habituality, where one would expect the aorist. Consequently, the function of the aorist is restricted to expressions of modality, such as predictions of future actions or necessity (see Bulut 2016).

The formal and functional vagueness, and the fact that the distribution of both present types is blurred (one may find present where one expects aorist, and vice versa), may create the impression that either the two have merged into one, or that a focal present is still not fully developed. This may be due to the language contact situation, as the Iranian languages have just one morphologically simple present. The negative forms of present and aorist, on the other hand, are morphologically distinct, see, for instance, the negative forms of the first person singular of the focal present *ged-mir-äm* ‘I don’t go (now)’, vs. the aorist *ged-mäm/ged-män-äm* ‘I never go’, or 1st plural *ged-mir-ik* ‘we don’t go’, vs. *gedmä:k* ‘we never go’, see Table 7 below.

Table 7: Negative forms of the present and the aorist

	present	aorist		present	aorist
1 SG	<i>ged-mir-äm</i>	<i>ged-mäm/ged-män-äm</i>	1 PL	<i>ged-mir-ik</i>	<i>gedmä:k</i>
2 SG	<i>ged-mir-äy</i>	<i>ged-mäz-äy</i>	2 PL	<i>ged-mir-i:z</i>	<i>ged-mä-yi:z</i>
3 SG	<i>ged-mir(i)</i>	<i>ged-mäz</i>	3 PL	<i>ged-mir-lär</i>	<i>ged-mäz-lär</i>

A non-evidential perfect

Across the region, the perfect displays three morphologically different types. Perfect paradigms are based either on the old morpheme {-mXš}, on a mixed paradigm consisting of {-mXš} and a ‘new’ perfect marker which goes back to the gerund in {-yXb}, or solely on {-yXb}. The personal suffixes are of the pronominal type, while the 3rd persons may optionally combine with the copula {+DXr} <*durur*.²¹

The Qashqâ’î-varieties of the Safî Khânî and the Bayat (see Bulut 2016), and also Sonqorî have preserved the old paradigm in {-mXš} in all persons, see the examples from sample text 1, with the characteristic /-š/ > [-Ø], such as: *gelmî* ‘come.PRF.3SG’, *ilemi* ‘make.PRF.3SG’, and *vermi* ‘give.PRF.3SG’ (1/4). Other varieties have completely restructured their perfect. Paradigms on the basis of the new element {-yXb} are found in Altunköprü in Iraq, in Qal‘â Gözälđärâ in the eastern part of Bayâdestân (bilingual Kalhor Kurds), in the Province of Esfahan, and in the Begdilli variety of Qashqâ’î. More frequent are mixed paradigms with {-mXš} for 1st persons and {-Xp} for 2nd and 3rd persons (most varieties of Iraq), or {-mXš} in the 1st and 2nd persons and {-Xp} for 3rd persons, as in Baku in North Azerbaijan and in Bayâdestân and other central varieties of Iran. The geographical distribution of the different perfect paradigms is extremely irregular, and often varies from one village to the next (see also Širaliev 1983: 115 on North Azerbaijani dialects). Moreover individual speakers may use different forms in parallel, producing third persons in {-yXb} and in {-mXš}, which could be interpreted as reflecting development in progress.

In contrast to the Turkish perfect in {-mIš}, the perfect occurring in the region we are concerned with (eastern Anatolia, western Iran, Azerbaijan and Iraq) is strictly resultative and has no marked inferential or evidential qualities. Inferential or evidential connotations are expressed by the evidential/inferential copula in {-ImIš}, as in *yatmiš-imis* ‘she had obviously gone to sleep’ (Khalaj of Bayâdestân);

²¹ According to Bodrogligli (1968: 30), the complex forms in {-yIBDIr} may have been inspired by the model of the Persian perfect, which—in contrast to the Turkic zero-morpheme—has an explicit morphological element to mark the 3rd person singular on the surface. Thus, the complex form *gälübür* ‘has come’ in the glossary of the Esfahan anonymous, for instance, would be a structural copy of the Persian *āmade ast* (come. PST.PTCP + copula 3SG).

see also a contracted form in sample text 1: *qejr-imı* (<*qeýrmiš-imıš*) ‘has founded (as they say)’ (1/2), and *mayoly_imı* (1/3) ‘was a Mongol (as I infer)’.

The imperative of the second person singular in {-ginen}

As in most Turkic languages, the bare verbal stem functions as the imperative of the 2nd person singular. Combinations with an additional suffix {-ginen}, which appear across the dialects of the area under consideration, express a somewhat milder or more polite form of request or demand, see example no. (4) below.

Potential and impossibility: {-abil-}, {-abilme-}

As is characteristic of the area, the potential verb form {-abil-} has the negation {-abilme-} (in contrast to {-y)AmA in Turkish of Turkey), e. g. *gäl-äbil-mäm* ‘I will not be able to come.’

The necessitative in {-AsI}

The necessitative in {-AsI} combines with suffixes of the copula (or: the so-called pronominal conjugation) to form predicates. Its predecessor is the so-called future in {-asi/-äsi}, which appears in a wider range of syntactic functions in Old Anatolian Turkic, see Mansuroğlu (1959: 173–174), and Bulut (2016). The modal character of this element (‘to have to do something’) may be re-enforced by copied modal adverbs, such as the Persian *hatman* ‘really, absolutely, by all means’, as in *sän hatman gid-äsi-yäj* ‘you absolutely have to go’ (Safî Khânî/Il-e Qashqâ’î). The entity is present in the Qashqâ’î varieties of the Saffî Khânî, Begdili, Bayat and Amaleh in the extreme Southwest of the area; while it seems to be missing (or has been replaced by the necessitative in {-mAlI})²² in the varieties of central Iran and Iraq, it re-surfaces in the northwest, in Azerbaijani proper (see Širaliyev 1983: 135).

Table 8: Necessitative mood

(1) Safî Khânî		Amaleh according to Jarring (1943–1944), contracted forms with *	North Azerbaijani dialects according to Širaliyev (1983: 135)
<i>män</i>	<i>hatman</i> <i>gid-äsi-yäm</i>	<i>gedäsim*</i>	<i>daniš-asi-yäm</i>
<i>sän</i>	<i>hatman</i> <i>gid-äsi-yäj</i>	<i>gedäsij*</i>	<i>daniš-ast-san</i>
<i>o</i>	<i>hatman</i> <i>gid-äsi</i>	<i>ged-äsi</i>	<i>daniš-ast-di</i>
<i>bız</i>	<i>hatman</i> <i>gid-äsi-yäg</i>	<i>ged-äsi-yäg</i>	<i>daniš-ast-yuy</i>
<i>sæn</i>	<i>hatman</i> <i>gid-äsi-yäniz</i>	<i>ged-äsi-yäniz</i>	<i>daniš-asi-suz</i>
<i>o^Ular</i>	<i>hatman</i> <i>gid-äsi-lär</i>	<i>ged-äsi-lär</i>	<i>daniš-asi-di (lär)</i>

²² Whether {-AsI} appears in complementary distribution with the necessitative in {-mAlI} is a question for further research.

2.4. Morpho-syntactic elements

2.4.1. The question particle {mI}

In the entire area the question particle {mI} does not appear. Questions are marked by a copied intonation pattern, the so-called drawl, which consists of a rising intonation and lengthening of the clause-final syllable; see also §2.5.2 below.

2.4.2. Copied morpho-syntactic units

2.4.2.1. Comparative in {+tar} and superlative {+tarin}

See under §2.3.1.4.

2.4.2.2. The specifiers in {+AkA} or {+i}

{+AkA} is a global copy of the Kurdish or, originally, Gorani marker of specificity (“definite article”). It appears across Turkic varieties in contact with these Iranian languages, such as, for instance, Iraq-Turkic, Sonqorî, and varieties of Qashqâ’î. Attached directly to nominal stems, it precedes possessive and case suffixes and adopts palatal harmony, see examples such as *dev-äkä* ‘the (aforsaid) demon’ (Iraq), and *şär-äkä-si-ni* (poem-SPEC-POSS.3SG-ACC) ‘this certain/aforsaid poem of his’ (Sonqorî). The unit also marks the head of restrictive relative clauses following Indo-European patterns. In this function, {+AkA} is competing with the Persian *yâ-ye ešāret* in {+i}, copies of which exist in most Iran-Turkic varieties (see below, §2.5.2.1).

2.4.2.3. Copies of Iranian enclitic personal pronouns

Global copies of Iranian enclitic personal pronouns

In southern varieties of Iraq-Turkic global copies of Iranian clitics—attached to the verb after the personal ending—represent the dative/benefactory case or the direct object; see Bulut, this volume, chapter 3.5, §2.4.4.2.

Similarly, Sonqorî uses the Iranian clitic of the 3rd person singular {-iš} to denote the direct object, as in: *apar-a-ši* (take-OPT.3SG-CLIT.PRON.3SG) ‘to take her away.’

Functional copies of Iranian enclitic personal pronouns

With regard to structure, Turkic possessive markers are similar to Iranian personal clitics. The main difference is functional: Turkic possessive suffixes are restricted to expressions of possession, while Iranian personal clitics denote the oblique case of the personal pronoun, e. g. genitive, directive/dative, and accu-

sative. Paraphrases for ‘to need’ (both in Iraq-Turkic and Iran-Turkic) and for ‘to own’ (occurring solely in Iran-Turkic) imply that in the given language contact situation, Turkic possessive suffixes have copied additional functional properties of Iranian clitics and may thus also represent the dative or benefactory case (see below).

Paraphrasing ‘to need’

To paraphrase ‘to need’, many Turkic languages use a combination of the copied Arabic noun *lâzîm* or the Turkic *gerek* ‘necessary’ and the personal pronoun in the dative/benefactory case, compare Standard Turkish *bana lâzîm* (me.DAT necessary (is)) ‘I need’.

In a different construction used across Iraq-Turkic and the central varieties of Iran-Turkic, the Turkic possessive suffix may function as a morphological marker indicating the benefactory case of the personal pronoun, see examples such as *lâzîm-im* (*dir*) (necessary-POSS.1SG (is)) ‘I need’, literally ‘necessary for me (is)’ (Iraq), or *lazîm-miz* (necessary-POSS.1PL (is)) ‘we need’ (Bayâdestân). The underlying structure implies that Turkic possessive suffixes have copied the functional property of Iranian clitics to express oblique/benefactory case (see also Bulut, this volume, chapter 3.5, §2.4.4.2).

Paraphrases for ‘to have’

Turkic languages do not possess a full verb to express ‘to possess, have’. Instead they use a construction based on the existence adjectives *var* ‘existent’ or *yox* ‘not existent’, while marking the possessor by a possessive suffix on the possessum: {NOUN.POSS + *var/yox*}; see sample text (2/2) *sâhib-xâne-miz var-idi* (landlord.POSS.1PL existent-COP.PST.3SG), literally: ‘Our landlord existed’, that is: ‘we had a landlord’.

Iran-Turkic also displays an alternative construction where reference to the possessor is encoded in a possessive suffix on the existence adjectives *var* or *yox*; see sample text no. (2/3) *errebât var-i* (relation existent-POSS.3SG), literally ‘his contact (is) existent’, that is: ‘he has contact’. As a rule, the latter construction seems to be limited to Turkic varieties spoken in Iran; sporadically it also occurs in the border region of Iraq around Khanaqin and Mandali. One may assume that again Iranian languages provided the structural model; so, for instance, Mackenzie (1981: 102) describes a similar paraphrase of ‘to have’ in the Kurdish dialects of Suleimaniye and Arbil. Kurmanji, on the other hand, displays the first type where the possessor is marked by a possessive suffix on the possessum, i. e. a structure corresponding to that of Standard Turkish. Across Iran-Turkic varieties speakers generally use both paraphrases for ‘to have’ in parallel.

Given the extended functional properties which Turkic possessive suffixes copied from Iranian clitics (see above), constructions such as *var-im* (existent.POSS.1SG) ‘I have’ or *yox-ey* (non-existent.POSS.2SG) ‘You don’t have’ may be

interpreted as a variety of the *mihi est* pattern: ‘X is existent for me’ ≈ ‘X belongs to me; I have/possess X’—with the Turkic possessive suffix representing the dative/benefactory case.

2.4.2.4. *The dative/directive in postverbal position: A copied word order property*

The Turkic dative case in {-y}A} may function as benefactory case, or as a directive case designating the aim to which an action expressed in the verb is directed. Across the region (Azerbaijan, Anatolia, Iraq, Iran), the dative often occurs in postverbal position.²³ This is probably due to the influence of Iranian languages such as spoken Persian or Kurdish, where postverbal position of the directive after verbs expressing motion is the rule (compare also Bulut, this volume, chapter 3.5, §2.4.5.1); see also the following example from sample text (1/1): *vä bu gälmi Songore* ‘and this (guy) came to Sonqor.DAT.’

2.4.2.5. *Turkic postpositions vs. copied Iranian prepositional phrases*

Although Iran-Turkic has preserved most of the Turkic inventory of postpositions, copied prepositional phrases after Iranian models are extremely frequent in varieties with strong Iranian influences, such as, for instance, Sonqorî. In sample text (1/1) both types of adposition appear in the same sentence, see *râje* ‘be târîh ‘about the history’, with the complex Persian preposition *râje* ‘be ‘concerning’, vs. *täpä üstündä* ‘on a hill’, with the Turkic secondary postposition *üst-iün-dä* (< surface/top-POSS.3SG-LOC), literally: ‘hill-on its top’; more examples of copied prepositional complexes in the same text are: *mærbud be* ‘pertaining to’, *az* ‘from, of’, both in (1/3), or *hattâ* ‘until, up to’ in (1/5).

2.4.2.6. *Copies of ezâfe constructions*

Across Iran-Turkic varieties, copies of the Iranian *ezâfe*, a unit which originally connects substantives, adjectives or pronouns to the head of a noun phrase, are used as a productive pattern. The copied *ezâfe* constructions, though, are restricted to combinations of (substantive-EZ + adjective) or (substantive-EZ + substantive), mostly connecting Iranian lexical elements. Again the frequency of *ezâfe*-constructions indicates the degree of Iranicization, as the following copies of Persian noun phrases from Sonqorî demonstrate: *târîh-e Songor* (history-EZ Sonqor) ‘the history of Sonqor’ in sample text (1/1), *mahall-e dafn-e mâlik-i* (place-EZ burial-EZ king-SPEC) ‘the place of burial of the king’ (1/2), or *u taraf-e Songor* (this side-EZ

²³ Note that Turkic languages may make use of similar deviations from the unmarked SOV word order to put a certain element in focus, which is not the case in constructions with the postverbal dative in Turkic of Iran and Iraq.

Sonqor) ‘this side of Sonqor’ (1/5), where the Persian *ezâfe* connects substantives; and *afsar-ân-e Mağol* (soldiers-PL-EZ Mongol) ‘the Mongol soldiers’ (1/3, with a Persian plural marker on the first element), or *zabân-e Torkî* (language-EZ Turkic) ‘the Turkic language’ (1/4), where the *ezâfe* links an adjective to a substantive. For Iraq Turkic, see Bulut, this volume, chapter 3.5, §2.4.5.2, where the usage of *ezâfe* constructions is rather limited.

2.5. Syntax

Turkic languages mainly rely on non-finite verb forms to express the syntactic equivalents of Indo-European dependent clauses. In principle *adverbial action clauses* are based on gerunds, while nominalized verb forms such as verbal nouns and participles denote *agent clauses* (relative clauses) or *nominalized action clauses* (such as, for instance, complement clauses).²⁴ In Iran-Turkic, and, to a lesser extent, also in Iraq-Turkic and in the Turkic varieties spoken in Azerbaijan proper, these principles of Turkic syntax are largely abandoned. In irregular distribution one comes across a very limited inventory of surviving gerunds and nominalized subordinators, examples of which are given under §8.1 below.

2.5.1. The Turkic principle of subordination: Non-finite subordinators

The following paragraphs 2.5.1.1 to 2.5.1.2 present examples of surviving Turkic non-finite subordinators which are rather unusual in Iran-Turkic.²⁵

2.5.1.1. Gerunds

Temporal clauses based on the complex gerunds in {-y)AndA} and {-y)AndAn} Characteristic of the region (including Iran, Iraq, eastern Anatolia and Azerbaijan) are complex gerunds in {-y)AndA} and {-y)AndAn}, combinations of the present participle in {-y)An} with the case suffixes of the locative or ablative (+ postposition *sonra*). The complex gerund in {-y)AndA} encodes a notion of simultaneousness (see ex. 1) or anteriority (ex. 2) of the action expressed in the temporal clause:

²⁴ See Johanson (1990: 199–200).

²⁵ With shorter speech samples, one may get the impression that non-finite subjunctions are completely extinct across Iran-Turkic. Most of the examples quoted here originate from Bayâdestân in central Iran, where the present author conducted a large-scale field study (see Bulut: forthcoming).

(1) Kerkuk

asmane attanda obüri daši qaldırı yerdän...
 sky.DAT throw.GER other.ACC stone.ACC takesup.PRS.3SG ground.ABL
 ‘While throwing (the first stone) up to the sky, he lifts up the other one
 from the ground, ...’

(2) Bayâdestân

čöldän gäländä bilälärä nahar âmad' elärdik
 outskirts.ABL come.GER them lunch ready make.AORII.1PL
 ‘When they came back from (the fields) outside the village, we would
 prepare lunch for them.’

The complex gerund in {-yAndAn} stresses anteriority of the action expressed in the gerundial phrase, as in example (3) below:

(3) Bayâdestân

xarman eliyännän sora o qururdu; quriyännän
 threshing floor come.GER POSP that dry.AOR.PST.3SG dry.GER
sora o ökiz aparardilar ...
 POSP that oxen bring.AOR.PST.3PL
 ‘After they had piled it (=the wheat) up on the threshing floor, it would
 dry; after it had dried, they/one would bring the oxen ...’

Iraq-Turkic uses complex gerunds based on the verbal nouns in {-DIK} or {-mEK} in a similar function (see Bulut, this volume, chapter 3.5, §2.5.1.1). In the Sâfi Khânî variety of Qashqâ’î, the verbal noun in {-DIK} combines with the locative in {+DA} and the ablative in {+DAn} to form complex temporal gerunds, as in *köc-dig-dä* (migrate.VN.LOC) ‘while migrating’, and *är-dig-dän* (reach.VN.ABL) ‘after having reached’ (see Bulut: 2016).

Other relics of gerundial syntax

While Iraq-Turkic still has preserved a rich inventory of gerunds²⁶, in most Turkic varieties of Iran gerunds are nearly completely extinct. Among the few exceptions to this rule are Khalaj proper (see Doerfer 1988: 130–145), and the varieties of the Bayat and Khalaj of Bayâdestân in central Iran. In example (4) below, the gerund in {-yIp} functions as a non-modifying terminal gerund (see Johanson 1990), reflecting a temporal sequence of the semantic contents of gerundial phrase and main clause, with different agents of main clause and gerundial phrase in the first sequence:

²⁶ Such as, for instance, the gerunds in {-yE}, {-yIp}, {-EnI (< -EII)}, {-yIncE}, {aorist+ -kEn}, and {-dIKčE}; see also Bulut, this volume, chapter 3.5, §2.1.3.1.

(4) Bayâdestân

*Ušayiy olar, män gälip sänä
 child.POSS.2SG be.AOR.3SG I come.GER you.DAT
 yetišäbilmäm. Sän gälip burda qalginän!
 reach.POSS.AOR.1SG You come.GER here stay.IMP*
 ‘When your child is born, I won’t be able to come and wait on you. So come and stay here!’

In double verb constructions the gerund modifies the action expressed in the finite verb, as, for instance, in *tutup gätir-* (< hold.GER bring) ‘to bring along’, or *götürüp ged-* (< take-away.GER go) ‘take away/take to another place’; see also *durub gäl-* (< get_up.GER come) in example (5) below:

(5) Khalaj of Bayâdestân

*Sän Tährannan durub gälmišäy.
 you Tehran.ABL get_up.GER come.PRF.2SG
 ‘You have come all the way from Tehran.’*

2.5.1.2. *Verbal nouns and participles**The verbal noun in {-mAK}*

The verbal noun in {-mAK} combines with possessive and case suffixes (similar to Turkish {-mA}) to form complements of verbs indicating motion/direction/beginning; cf. the verbal noun in the dative after the aspect verb *şoru* ‘elä- ‘to begin’²⁷ in example no. (6), and ({-mAK} + Poss + DAT) in example (7) below:

(6) Bayâdestân

*Oylan şoru ‘elärdi gedmägä o qızın
 boy begin.AOR.PST.3SG go.VN.DAT that girl.GEN
 ävinä.
 house.POSS.3SG.DAT
 ‘So the boy would begin to visit the house of the girl.’*

(7) Bayâdestân

*Gidäk qızıy görümäginä.
 go.OPT.1PL girl.GEN see.VN.POSS.3SG.DAT
 ‘Let’s go to see the girl!’*

²⁷ See also Bulut, this volume, chapter 3.5, §2.4.5.1, example (7).

The verbal noun in {-DIK}

In some rare instances, the verbal noun in {-DIK} combined with possessive suffixes as agent markers and additional case suffixes forms absolute relative clauses, as in example (8) below:

- (8) Bayâdestân

<i>apardoyumütüy</i>	<i>da</i>	<i>ordä</i>	<i>yârsı</i>
bring.VN.POSS.1PL.GEN	too	there	half.POSS.3SG
<i>iyärđig</i>	<i>gälärđig.</i>		
eat.AOR.PST.1PL	come.AOR.PST.1PL		

‘Normally we would eat half of what we brought along over there before we returned home.’

The so-called present participle in {-y)En}

While Turkish employs the so-called present participle in {-y)An} primarily as a non-finite subordinator in constructions corresponding to Indo-European relative clauses (agent clauses), the unit functions as a nomen actoris in the Turkic varieties across the region, as in *işläyän* ‘worker’.

The participle is normally not attributed to a head of a relative clause. Examples such as *gälin gälän gecä* (bride come.PTCP night) ‘the night the bride arrived’ or *hava xoš olan günnär* (weather fine be.PTCP days) ‘days when the weather was fine’ (both examples from Bayâdestân) are exceptional; note that the head in both cases semantically refers to time.

Thirdly, the participle occurs in a special type of headless/absolute relative clause. In such constructions both participle and finite verb form derive from the same stem, the agents of subordinator and verb being identical, e. g. *qačan qačar imiš* (flee.PTCP flee.AOR.3SG + EV.COP) ‘and (they relate that) whoever was able to flee run away’, or *aylayan aylasun* (weep.PTCP weep.IMP.3SG) ‘who(ever) is weeping shall go on/who is able to weep shall do so.’.

Fourthly, the participle in absolute position may display its verbal potential to take objects, as in *säni veräniy övi ábâdan ossun!* (you.ACC give.PTCP.GEN house.POSS.3SG prospering be.IMP.3SG) ‘May the house of the one who gave you (into marriage) prosper!’

2.5.2. Syntactic patterns imitating Indo-European strategies of clause combining

Most Turkic varieties of the region have almost completely replaced the Turkic syntactic patterns of infinitization/nominalization. Like in the corresponding Indo-European or Iranian models, the dependent clause is based on a finite verb form. The finite verb may be connected to the main clause by a conjunction, which, at the same time, expresses the semantic relation between main clause

and dependent clause. Consequently, the Turkic varieties of the region exhibit a reduced inventory of non-finite subordinators, such as gerunds, verbal nouns and participles. The far-reaching Iranicization of syntax is strongest in the varieties spoken in Iran proper, and nowadays less prominent in the varieties of the peripheries.

Structural copying in syntax is often accompanied by phonological copying. Copied syntactic structures may also imitate supra-segmental intonation patterns of the model. The copied sentence melody or the respective patterns of pitch are part of the process of syntactic copying. In questions and co-ordinated clauses formed according to Iranian models, the tone often rises towards the end; additionally, the vowel of the last syllable may be lengthened, with a markedly rising and afterwards shortly falling tone. Like, for instance, in spoken Persian this so-called drawl—the rising tone at the end of a clause—signals that the sentence will go on, or that a semantically connected (main-)clause will follow. See sample text no. (1/2,3), where a drawl is indicated by ↗.

2.5.2.1. Indo-European relative constructions

The structure of Indo-European relative constructions is fundamentally different from that of Turkic syntactic patterns. While Turkic (attributive) relative clauses are based on a non-finite subordinator preceding the head, Indo-European post-positive relative clauses exhibit a finite verb form. A conjunction indicates the semantic relation between main clause and dependent clause.

The underlying pattern of copied restrictive relative clauses in the Turkic varieties of the area consists of {(DEM) + HEAD (+SPEC) + CONJ}, see examples (9) to (11) below; the head of the relative clause may be introduced by the cataphoric pronoun *o* or *ol* ‘that’, as in example (9) below, or anaphoric *bu*, as in example no. (10); see also sample text (2/6).

(9) Bayâdestân

<i>O</i>	<i>qizi</i>	<i>ki</i>	<i>istillär</i>	<i>alallar</i>
that	girl.ACC	CONJ	want.PRS.3PL	take.AOR.3PL
<i>oğlannarina.</i>				

son.POSS.3PL.DAT

‘They will chose the girl they like best for their sons.’

(10) Khalaj of Bayâdestân

<i>Andan</i>	<i>bu</i>	<i>kurelär</i>	<i>ki</i>	<i>qaziblär</i>	<i>her</i>
moreover	these	shaft.PL	CONJ	dig.PRF.3PL	every
<i>biri</i>	<i>älli</i>	<i>metirdir.</i>			

one.POSS.3SG fifty metre.COP.3SG

‘Moreover, every single one of these (aforesaid) side-canals they have dug is 50 metres long.’

- (11) Bayâdestân
ušaxlar ki taze 'amele²⁸ gällär Fârsî
 child.PL CONJ new world.DAT come.PRS.3PL Persian
danişirlär.
 speak.PRS.3PL
 ‘The children who are born now speak Persian.’

Eventually, copies of Iranian markers of specificity may be attached to the head, such as, for instance, the Persian *yâ-ye ešâret* in examples (12) from Sonqor and (13) from the Khalaj of Bayâdestân, or the Kurdish clitic {+Aka}, as in example (14) from the Abivardi variety of Shiraz (after Tuna 1984: 227):

- (12) *Bu mahall-e dafn-e malik-i ri*
 this place-EZ burial-EZ king-SPEC COP.3SG
ke Songor qeyrimi.
 CONJ sonqor found.INF.PRF.3SG
 ‘This is the place of burial of the king who (they say) has founded the town of Sonqor.’
- (13) *Be Xoda, o Xoda-yi kê mäni yaratub,*
 with God that God.SPEC CONJ me create.PRF.3SG
bu yuxardan bu üč tane
 this further_up.ABL this three QUANT
qurt hämçin mäni či kár elärdilär?
 Wolf also me what do.AOR.PST.3PL
 ‘By God, by the Lord who has created me, what would these three wolves up there do to me?’
- (14) Abivardi²⁹
kitab-eku ke aldem
 book.SPEC CONJ I bought
 ‘The book I bought’

2.5.2.2. Temporal clauses

A pattern with identical surface structure {(DEM) + NOUN (+SPEC) + CONJ} forms temporal clauses. It also owes its existence to copies of a model that is widespread across spoken Persian and Kurdish (for Iraq Turkic see Bulut, this volume, chapter 3.5, §2.3.2.2); see example (15) from sample text (2/1) and example (16) below, which exhibit the same surface structure as the restrictive relative clauses discussed in §2.5.2.1:

²⁸ Metathesis of ‘âlem ‘world’.

²⁹ The example is presented in Tuna’s transcription.

- (15) *Män özüm ke därs oxibardim dabirestândan,*
 I myself CONJ study.AOR.PST.1SG high_school.ABL
Noubarânda bir yerdä oturardıq.
 noubarân.LOC some place live.AOR.PST.1PL
 ‘When I was studying at high school, we were living in a place in
 Noubarân.’
- (16) Bayâdestân
o šoxm ke elämišdik äkärdik.
 that plough CONJ do.PLUP.1PL sow.IMP.1PL
 ‘After we had ploughed (the fields), we used to sow.’

2.5.2.3. Complement clauses

Turkic complement clauses may rely on non-finite subordinators (verbal nouns) as the direct object of *verba dicendi et sentiendi*. Across the region this construction is replaced by copies of Iranian patterns (for Turkic in Iraq, see Bulut, this volume, chapter 3.5, §2.3.2.3): The postpositive complement clause is based on a finite verb form and is linked to the main clause by the copied conjunction *ki/ke*, see sample text (2/6), and examples (17)–(19) below. The same construction is applied to render indirect speech, as in example (19).

- (17) Kalhor of Bayâdestân
Görällä’ ke dädäm huşa gälär
 see.AOR.3PL CONJ father.POSS.1SG consciousness come.AOR.3SG
 ‘They saw that my father recovered consciousness.’
- (18) *Bilsinler ki mänim naharım yanibdi da*
 know.IMP.3PL CONJ my lunch.POSS.1SG burn.PRF.3SG also
 ‘They shall know that my lunch has burned.’
- (19) Kalhor of Bayâdestân
Dedim ki ilan čox daha varde.
 say.PST.1SG CONJ snake many more existent.COP.PRS.3SG
 ‘I said that there were many more snakes.’

2.5.2.4. Periphrastic expressions of modality

Turkic languages possess a diverse inventory of finite verb forms to express various moods, such as imperative, voluntative, necessitative or conditional. Moreover, nominalized verb forms/non-finite subordinators, such as the so-called future participle, the necessitative {-mAll}, or combinations of an impersonal modal expression (such as *gerek* or the copied *lázim* ‘necessary’) and the verbal nouns in {-mA} or {-mAk} etc. can express distinctive shades of modality. In Iranian languages, on

the other hand, the subjunctive is the sole category that indicates mood or marks dependent verbs. In situations of intensive contact with Iranian languages (or other Indo-European languages, such as, for instance, Greek), many Turkic varieties have restructured their means of expressing modality, with the Turkic optative-subjunctive or the imperative spreading into positions of equivalence held by the multi-functional Indo-European subjunctive (Soper 1996: 284–285; Bulut 2000, and 2009: 67–68) — a development which is already reflected in Old Anatolian Turkic texts.

The intentional ‘to want to’

The Turkic expression for ‘to want to do something’ consists of a lexical verb in the form of a verbal noun that functions syntactically as a complement of the modal verb ‘to want’. Turkish, for instance, would render ‘I want to go’ as *gitmek istiyorum* {go.VN want.PRS.1SG}. Turkic varieties across the region, and, especially Iran-Turkic, use a periphrastic construction consisting of *iste-* ‘to want’ and a dependent lexical verb in the optative-subjunctive or the imperative, such as, for instance, *istiräm gidäm* {want.PRS.1SG go.OPT.1SG}, and *istiräm gitsin* {want.PRS.1SG go.IMP.3SG} ‘I want him/her to go’. As the three translations of the sentence ‘I want to go’ below demonstrate, the underlying model is the Iranian periphrastic modal construction, the order of constituents being modal verb—lexical verb, while the optative in Iran- or Iraq-Turkic corresponds to the subjunctive of the dependent verb in the Iranian model.

Turkish		Turkic varieties	Iran/Iraq	Modern Persian
<i>Gitmek</i>	<i>istiyorum</i>	<i>İstiräm</i>	<i>gidäm</i>	<i>Mi-xwâham</i>
go.VN	want.	want.	go.OPT.1SG	DUR-want.
PRS.1SG		PRS.1SG		PRS.1SG

Possibility

In the Turkic varieties of the area the so-called (im)possibility form (‘to be (not) able to’) exhibits a strongly grammaticalized combination of lexical verb and the former modal verb (*bil-* ‘to know, to be able to’)³⁰, e. g.

- (20) Kalhor of Bayâdestân

Gördüm daha gidäbilmir.
see.PST.1SG more go.IMPOSS.PRS.3SG
'I realized that it couldn't walk anymore.'

³⁰ To express (im-)possibility or (in-)ability, the Turkic varieties of the region (Azerbaijan, Iraq, Iran and some regions in eastern Anatolia) may employ the deverbal verb suffix {-Ebil-} attached to the stem of the lexical verb. As for negated forms, {-Ebilmi-}, as in (2: 30) *yiy-abilmæz* {collect-IMPOSS.AOR.3SG} ‘he will not be able to collect’, or {-EmE}, as in Turkish of Turkey, appear in irregular distribution, or, in some regions, even in parallel.

The notion of (im-)possibility or (in-)ability may also be periphrased by using the Turkic full verb *başar-* ‘to succeed, can, be able to’ as an auxiliary, which combines with the lexical verb in the subjunctive/optative, see example (21) from Sonqorî:

(21) Sonqor

<i>başarmazmı(š)</i>	<i>qaza</i>	<i>qeyrā</i>
succeed.AOR.NEG.INF.3SG	food	prepare.OPT.3SG
'She is (as they say) not able to/cannot prepare food.'		

The model for this periphrastic construction is in all probability the Iranian structure relying on the modal auxiliary ‘can’ and the lexical verb in the subjunctive, as, for instance in the Persian *mî-tavân-am be-gu-y-am* or the Kurdish *di-kar-im be-ž-im* {DUR-can.PRS-1SG + SUBJ-say.1SG} ‘I can say’, literally: ‘I can (that) I say’.

‘to tolerate, let’

Structurally similar is the expression for ‘let someone do something’, which is based on the Turkic verbs *biraq-* ‘let’- or *qo(y)-* ‘put’ as auxiliaries, combining with the lexical verb in the optative or imperative. This periphrastic modal construction clearly goes back to a calque of the Iranian/Persian compound expression for ‘to let someone do something’. As in the case quoted above, in Persian a former full verb (*gozâštan ~ godârdan ~ gozârdan*, present stem: *godâr ~ gozâr* ‘to put, let, leave, permit, tolerate, suffer’) with reduced lexical content takes up the function of an auxiliary, meaning ‘let, permit’. It combines with the lexical verb in the subjunctive, as in *bezâr³¹ bexâbam* {let.IMP.2SG sleep.SUBJ.1SG} ‘let me sleep!’. Both in the Persian model and in the Turkic copy, the agents of the auxiliary and lexical verbs may differ, see examples (22) and (23) below:

(22) Kerkuk

<i>Qoy</i>	<i>zad</i>	<i>sänä</i>	<i>deyim.</i>
put.IMP.2SG	thing	you.DAT	say.OPT.1SG
'Let me tell you something.'			

(23) Bayâdestân

<i>Qoymadi</i>	<i>qačsin.</i>
put.NEG.PST.3SG	get_away.IMP.3SG
'He didn't let him get away.'	

³¹ Note that in colloquial Persian, the stem-initial /g-/ is deleted; *bezâr* thus is a short/contracted form of the imperative of the 2nd person singular *begozâr*.

Impersonal modal expressions

Impersonal modal expressions such as *gäräk* or *lâzîm* ‘necessary’ also combine with the lexical verb in the optative or imperative, which hold a position of equivalence to the Iranian subjunctive (see Bulut, this volume, chapter 3.5, §2.5.2.4, and Bulut 2000). The Iran-Turkic varieties under consideration quite regularly display the optative in corresponding constructions, see, for instance, e. g.

- (24) Esfahan

<i>härkäs</i>	<i>šärbätinän</i>	<i>gäräk</i>	<i>dâmâd</i>	<i>ičä</i>
everyone	sorbat.POSS.3SG.ABL	necessary	bridegroom	drink.OPT.3SG
‘the bridegroom has to drink from everyone’s lemonade ...’				

Instead of genuine Turkic expressions such as *gäräk*, global copies of impersonal Persian modal expressions may also appear, as, for instance, *na-bâdâ* ‘it shall not be; may God forbid that ...’ in example (25) from Sonqor:

- (25) Sonqor

<i>Na-bâdâ</i>	<i>äl</i>	<i>gälä</i>	<i>bunu</i>	<i>apara.</i>
it shall not be	al	come.OPT.3SG	her.ACC	take away.OPT.3SG
‘It shall not be that the Al comes to take her away.’				

Expressions of purpose and intention

The notion of finality (intention or purpose/aim to which the action is directed) can be expressed the Turkic way by a non-finite subordinator (the verbal noun in {-mAK}) combined with the postposition *için*, see Bulut, this volume, chapter 3.5, §2.5.2.4. More frequently used, though, is a construction imitating Iranian models, where the non-finite verb form is replaced by a finite verb in the subjunctive/optative as in example (26), or in the imperative.

- (26) Khalaj of Bayâdestân

<i>yari gecä</i>	<i>Al</i>	<i>gälär</i>	<i>munuy</i>	<i>ušayın</i>	<i>apara</i>
midnight	Al	come.AOR.3SG	her	child.POSS.3SG.ACC	take_away.OPT.3SG
‘At midnight the Al comes to take her child away.’					

2.5.2.5. Conditional clauses³²

Turkic conditionals establish a separate category of mood explicitly marked by the morpheme {-sA}, which may combine with additional tense/aspect markers. The reduction of Turkic modal categories according to Iranian models based solely on the subjunctive (see above) has also affected Iran-Turkic expressions of conditionality. Thus, the optative (which has widely been identified with the Iranian sub-

³² For a more thorough discussion and more examples on Iran-Turkic conditional constructions see Bulut (2009).

junctive) now also appears in Turkic conditional clauses. As in the Iranian model, conditionality is no longer expressed by morphological means, but by the copied Iranian conjunction *ägär* ‘if’; see the conditional construction expressing potential in example (27) and irrealis (signalized by the optative past) in example (28) below:

- (27) Khalaj of Bayâdestân

*ägär čörägi tökäy toyuyä, (...) bäräkäti
if bread.ACC pour.OPT.2SG chicken.DAT blessings.POSS.3SG
gidär.
go.AOR.3SG*

‘If you give the bread to the chickens, [...] its *barakat*, its blessing will go.’

- (28) Bayâdestân

*Äyän onnan vurmiyaydim, hälnuz qani durmamišdi.
if it.ABL pour.OPTII.1SG still its blood stop.NEG.PLU.3SG*
‘If I had not made use of it (medicine), the blood would still be flowing.’

2.6. Lexicon

The Turkic core vocabulary is based mainly on Turkî, that is the South Oghuz or Central Oghuz (Azeri) lexicon. Characteristic of Iran-Turkic is a high proportion of Persian vocabulary, the ratio of which varies across the different varieties. At present most speakers of Turkî have at least a bi-lingual lexicon of Turkic-Persian synonyms at their disposal. Thus, it is difficult to say to what degree copies of Persian words are conventionalized in any particular variety, or which copies have outnumbered their Turkic counterparts. For a good survey of Iran-Turkic lexicon, see Hacaloğlu (1992).

Copies of the lexicon of smaller regional languages obviously pertain to older strata of the language. The Turkic tribes of Bayâdestân, for instance, have copied special terminology pertaining to agriculture and irrigation techniques from the neighbouring Tati populations. This implies that the nomadic tribes, who came into the area roughly two centuries ago, adopted techniques of cultivation and the related lexicon of their close neighbours.

No research exists on idiomatic expressions, many of which use the same semantic means across the different languages of the greater area.

Text 1 (Sonqor)

On the history of Sonqor

- (1/1) ?æmma: unʃi₁ k^he: radʒi bɪ ta:rɪx-e: Soŋcor
 but so CONJ about history Sonqor
 daniſſile:, bɪdʒe t^hep^hø_yssinen, bɪdʒe bu
 speak.AOR.3PL one hill.POSS one this
 gonbæd-mvñend saxtəman wa:r k^he: bo^o gunbæd
 dome-like building exist CONJ this dome
 Sældžuyılær zæmƿni: 'bænnpsinne_dej.
 Saljuq.PL time.POSS.3SG building.POSS.3SG.ABL.COP.3SG
 'Yet, concerning how people talk about the history of Sonqor, there is a dome-like building on the top of a hill, and this dome is from the Saljuqian period.'
- (1/2) vε di'jelle: ɬ, mæħell-ε dæfn-e mp:lik^h-i₁ri
 and say.AOR.3PL place.EZ burial.EZ king.SPEC.COP.3SG
 k^he: Soŋgore qejrimi.
 CONJ Sonqor.ACC found.INF.PRF.3SG
 'And they say that it is the place of burial of the king who founded Sonqor.'
- (1/3) vε b'dʒe: dε mot^hæyid_dələ:³³ k^hi bu qabər-1 k^hi
 and some also convinced.COP.3PL CONJ this grave.SPEC CONJ
 wør_di: ɬ mærbud di: be Soŋcor bæk^h, k^he
 exist.COP.3SG pertaining.COP.3SG PREP Sonqor Beg CONJ
 æz sær-bvɔzɔn_ o æfsærɔn-e mayol_imi.
 PREP soldier.PL and officer.PL.EZ Mongol.INF-COP.3SG
 'And some are also convinced that this grave over there pertains to Sonqor Beg, who is said to have been of the soldiers or officers of the Mongols'.
- (1/4) vε bu gelmi Soŋgore, urdæ wæt^hæn iləmi
 and this come.PRF3.SG Sonqor.DAT there home make.PRF.3SG
 zendegi_ iləmi; vε bo zæbɔ:n-e t^hørk^hini:
 life make.PRF.3SG and this language.EZ Turkî.ACC

³³ From Persian (< Arabic) *motaqâ 'ed* ‘convinced’.

ræwndʒi vermi,³⁴ k^he æl?ɒn tɑ_ 'nedʒin³⁵ bɒ:yı yalmi³⁶.
 promote.PRF.3SG CONJ now till now remain.PRF.3SG
 ‘So he came to Sonqor and settled down and spent his life there; and he promoted the Turkic language that has survived to this day’.

- (1/5) vɛ lb: t^hæmpɔ:³⁷ hætt^hb jæk^h færſæx, ſiʃ k^hilometr o
 and not just till one parasang six kilometres that
 t^hæræf-ə Sonqor t^hæmɒm Kurdı soþet^h ille:.
 side.EZ Sonqor all Kurdish speak.AOR.3PL
 ‘And not even one parasang, six kilometres to the other side of Sonqor, they only speak Kurdish.’

Text 2 (Bayâdestân)

The wedding of the jinn

- (2/1) mɛn øzym k^he dærſ oxiſardim dævirisda:n,
 I myself CONJ lesson read.AOR.PST.1SG high_school.ABL
 nouwærande bir jerd' oturardeq.
 Noubarân.LOC one place.LOC live.AOR.PST.1PL
 ‘When I was studying at high school, at the *dabîrestân*, we were living in a place in (the town of) Noubarân.’³⁸
- (2/2) mostædʒir_ədig, s^hv:θəb-xvnamez var_ədə, o ʃox
 tenant.COP.PST.1PL landlord.POSS.1PL exist.PST.3SG he very
 ə dʒinna'-o v:ldan tæθreflerda³⁸.
 that jinn.PL-CONJ Al.ABL relate.AOR.PST.3SG
 ‘We had rented the place, and our landlord used to talk a lot about these demons and the [birth-demon] *Al*.’

³⁴ From Persian *ravāğ dâdan* ‘to spread, promote’.

³⁵ Contracted form of *indicin* ‘now.INSTR’

³⁶ From Persian *bâqî mândan* ‘to remain’.

³⁷ For Arabic *ka-mâ* ‘just, only’.

³⁸ For *ta'rîf elâ-* to describe, praise, relate etc.

- (2/3) mæsælæn deerdi: “bir dſæne noubærændeⁱ
 for_instance say.AOR.PST.3SG one QUANT Noubarân.LOC
 kʰiſe var, dʒinner_ənan ertəwat³⁹ varə.”;
 man exist jinn.PL.INSTR relation exist.POSS.3SG
 ‘So, for instance, he would say: “There is a man here in Noubarân who is
 in contact with the demons”;
- (2/4) in_kʰe: bir kʰæſən tojinda: bu dʒinner,
 CONJ one person.GEN wedding.POSS.3SG.LOC this jinn.PL
 derə, mæsælæŋ geliller, bu bızim
 say.PRS.3SG for instance come.PRS.3PL this our
 vðæmij zvdlærin, libɔ:ſin, zvð_
 man.GEN thing.PL.POSS.3SG.ACC clothes.POSS.3SG.ACC thing
 vparellej ge?illær.
 take_along.PRS.3PL wear.PRS.3PL
 ‘so that, as he said, at somebody’s wedding these demons would come
 and take this person’s things, his clothes along to wear them.’⁴⁰
- (2/5) annan o kʰiſine dʒinner, de?eller, tojlarinə,
 CONJ that guy.ACC jinn.PL say.AOR.3PL wedding.POSS.3PL.DAT
 de?eller o dʒinner_ənæn ræfd-vmæd var,
 say.PRS.3PL that jinn.PL.with contact exists
 dæwet elellær bylæsin.
 invitation make.AOR.3PL him.ACC
 ‘Now people say that the jinn invite this man who keeps contact with
 them to their weddings.’

³⁹ From Persian (< Arabic) *ertebâfl* ‘relation, contact’.

⁴⁰ This story is somehow related to the belief of the *ta’bia*, a twin everyone (or especially every woman?) has among the people of the jinn. If a girl gets married, her alter ego of the demons will hold a wedding on the same day. Therefore, she borrows her human counterpart’s wedding dress.

- (2/6) annan kʰiʃi görər kʰi:, bu qoŋʃosinən
then guy see.PRS.3SG CONJ this neighbor.POSS.3SG.GEN
qizeniň liba:su^j kʰi tʰojde
daughter.POSS.3SG.GEN dress.POSS.3SG CONJ wedding.LOC
g̊e:jm^lsde, o gæ:linj kʰi, dʒi:n kʰi gæ:lin
wear.PLUP.3SG that bride CONJ jinn CONJ bride
olvb, onən ne ejnēnde də.
become.PRF.3SG her also body.POSS.3SG.LOC COP.3SG
‘And there this man sees that the bride, the bride of the jinn, is wearing
the bridal dress that his neighbor’s daughter had been wearing at her
wedding.’

Abbreviations

ABL	Ablative	INSTR	Instrumental
ACC	Accusative	LOC	Locative
AOR	Aorist	NEG	Negation
CONJ	Conjunction	OPT	Optative
COP	Copula	PL	Plural
DAT	Dative	PLUP	Pluperfective
DEM	Demonstrative	POSP	Postposition
EZ	Ezafe	POSS	Possessive
GEN	Genitive	PRF	Perfective
GER	Gerund	PST	Past
IMP	Imperative	SG	Singular
IMPOSS	Impossible	SPEC	Specificity
INF	Inferential		

References

- Boeschoten, Hendrik. 1998. The speakers of the Turkic languages. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic Languages*, 1–15. London & New York: Routledge.
- Bulut, Christiane. 1999. Klassifikatorische Merkmale des Iraktürkischen. *Orientalia Suecana*, vol. 48. 5–27.
- Bulut, Christiane. 2000. Optative constructions in Iraqi-Turkman. In Aslı Göksel & Celia Kerslake (eds.), *Studies on Turkish and Turkic Languages. Proceedings of the IXth International Conference on Turkish Linguistics, Lincoln College, Oxford, 12–14 August, 1998*, 161–169. (Turcologica 46). Wiesbaden: Harrassowitz.
- Bulut, Christiane. 2003. Pronominal systems in the transitional varieties of the Turkic dialects in East Anatolia, Iraq and Western Iran. In Sumru Özsoy (ed.), *Studies in Turkish Linguistics: Proceedings of the Tenth International Conference in Turkish Linguistics, Bogaziçi University, 16–18 August, 2000*, 321–335. Istanbul: Bogaziçi Yayınları.

- Bulut, Christiane. 2009. Conditional clauses in Iran-Turkic. In Éva Á. Csató, Gunvald Ims, Joakim Parslow, Finn Thiesen & Emel Türker (eds.), *Turcological letters to Bernt Brendemoen*, 35–75. Oslo: The Institute for Comparative Research in Human Culture. Novus Press.
- Bulut, Christiane. 2016. Convergence and variation in the Turkic varieties of Iran: Examples from Qashqâ'i. In Éva Á. Csató, Lars Johanson, András Róna-Tas & Bo Utas (eds.), *Turks and Iranians: Interactions in Language and History*, 235–282. (Turcologica 105). Wiesbaden: Harrassowitz.
- Bulut, Christiane. Forthcoming. *On language and culture in the Turkic region of Bayâdestân/Iran*. Habilitation thesis, Mainz 2006.
- Islamov, M. I., Aghajev, Ä. G., Behbudov, S. M., Ahmadi, T. M., Mammadov, N. İj., Tagiyev, B. M. & Z. A. Hasiyev (eds.). 1990. *Azärbaycan Dilinin Dialektoloji Atlası* [Dialectological Atlas of the Azerbaijani language]. Baku: Älm.
- Doerfer, Gerhard. 1988. *Grammatik des Chaladsch*. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 2006. Irano-Turkic. In Lars Johanson & Christiane Bulut (eds.), *Turkic-Iranian Contact Areas*, 93–113. (Turcologica 62). Wiesbaden: Harrassowitz.
- Doerfer, Gerhard & Wolfram Hesche. 1989. *Südoghusische Materialien aus Afghanistan und Iran*. Wiesbaden: Harrassowitz. Hacıaloğlu, Recep Albayrak. 1992. *Azeri Türkçesi Dil Kılavuzu* [Azeri Turkish Language Guide]. Ankara: Türk Tarih Kurumu.
- Gabain, Annemarie von. 1941. *Altürkische Grammatik*. Wiesbaden: Harrassowitz.
- Hacıaloğlu, Recep Albayrak. 1992. *Azeri Türkçesi Dil Kılavuzu* [Azeri Turkic language guide]. Ankara: Türk Tarih Kurumu.
- Johanson, Lars. 1985. Die mittelaserbaidschanischen Transkriptionstexte der Universitätsbibliothek Uppsala. In *Beşinci Milletler Arası Türkoloji Kongresi. İstanbul 1985. Tebligler I. Türk Dili, Cilt 1*. 143–147.
- Johanson, Lars. 1990. Subjektlose Sätze im Türkischen. In Bernt Brendemoen (ed.), *Altaica Osloensis. Proceedings from the 32nd Meeting of the Permanent International Altaistic Conference*, 193–218. Oslo: Universitetsforlaget.
- Johanson, Lars. 1993. Code-copying in immigrant Turkish. In Guus Extra & Ludo Verhoeven (eds.), *Immigrant languages in Europe*, 197–221. Clevedon, Philadelphia & Adelaide: Multilingual Matters.
- Mann, Oskar. 1930. *Die Mundarten der Gûrân, besonders der Kändülâî, Auramânî und Bâdschälâî*, (ed. by Karl Hadank). Berlin.
- Mansuroğlu, Međut. 1959. Das Altosmanische. In *Philologiae Turcicae Fundamenta I*, 161–181. Wiesbaden: Steiner.
- Şiraliyev, M. S. 1983. *Dialekty i govory azerbajdžanskogo jazyka* [Dialect and speech of the Azerbaijani language]. Baku.
- Soper, John. 1996. *Loan Syntax in Turkic and Iranian*. (Rev. and ed. by Andras J.E. Bodrogliglieti). Bloomington, Indiana: Eurolingua.
- Tuna, Osman Nedim. 1984. Ebi Verdi: Iran'da bir Türk diyalekti [Ebi Verdi: A Turkic dialect in Iran]. *Türk Dili Arastırmaları Yıllığı, Belleten*, 215–245.
- Windfuhr, Gernot. 2006. Linguistic Sketch of Azeri. In *Encyclopædia Iranica* 8.4, 396–401. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/iran-vii7-turkic-languages> (accessed 15 February 2018).

4.3. Bakhtiari

Erik Anonby and Mortaza Taheri-Ardali

1. Introduction

The Iranic (Indo-European) language Bakhtiari is spoken by about one million people in the Zagros Mountains of south-western Iran. The main part of the Bakhtiari language area is divided among four provinces: Khuzestan, Chahar Mahal va Bakhtiari (=C&B), Lorestan and Esfahan. The largest cities with significant Bakhtiari-speaking populations are Masjed Soleymān and Izeh (Khuzestan), Shahr-e Kord (C&B), Dorud (Lorestan) and Aligudarz (Lorestan) (Anonby and Asadi 2014, hereafter AA 2014). The approximate location of the Bakhtiari language area is shown in Fig. 1.

The Bakhtiari ethnic confederation (B. *ayl*) has played a prominent role in the history of Iran (Digard 1988), and their nomadic livelihood is one of the iconic images of cultural diversity in the country. Aspects of Bakhtiari culture, such as the weaving of carpets and traditional clothing, including the distinctive indigo-and-white striped woolen vests (B. *čuqā*), are well-known inside and outside of Iran (Ittig 1988; Digard 1992).

Within the Iranic family, Bakhtiari has been classified as Southwestern, and more specifically as part of the Lori (or “Perside”) language group (Windfuhr 2009; MacKinnon 2011; Anonby 2017). The Lori group forms a continuum between Persian and the Kurdish language bloc to the north-west (Anonby 2003, 2016). Bakhtiari is geographically and structurally intermediate between Northern Lori and Southern Lori, although it is more similar to the latter (Amanollahi 1991; MacKinnon 2011; Anonby 2003, 2017). Within the Bakhtiari language, no clear dialect groups have been identified; speakers tend to refer to subvarieties of Bakhtiari in reference to urban centres, or major ethnic divisions (e.g., the *Haflang* vs. *Chārlang* ethnic moieties) (AA 2014: 17), but no linguistic evidence has been advanced to support these assessments, and an overall picture of the internal dialect situation remains unclear.

In the first published study relating to Bakhtiari, Mann (1910) made general remarks on the language and collected some texts as part of a larger work on the Lori language group. This was supplemented by a more substantial description of Bakhtiari by Zhukovskij, which was published posthumously in 1923 (Abdullaeva 2009), and recently re-edited and translated into Persian by M. Shafaghi and M. Dadres (Zhukovskij 1923/2017). Lorimer, whose (1922) work is the best-known study of the language, treated Bakhtiari phonology in a way that is both extensive and detailed, but not phonemically rigorous. He also prepared and published many Bakhtiari texts (Lorimer 1930, 1954, 1955, 1963, 1964; Lorimer and Lorimer

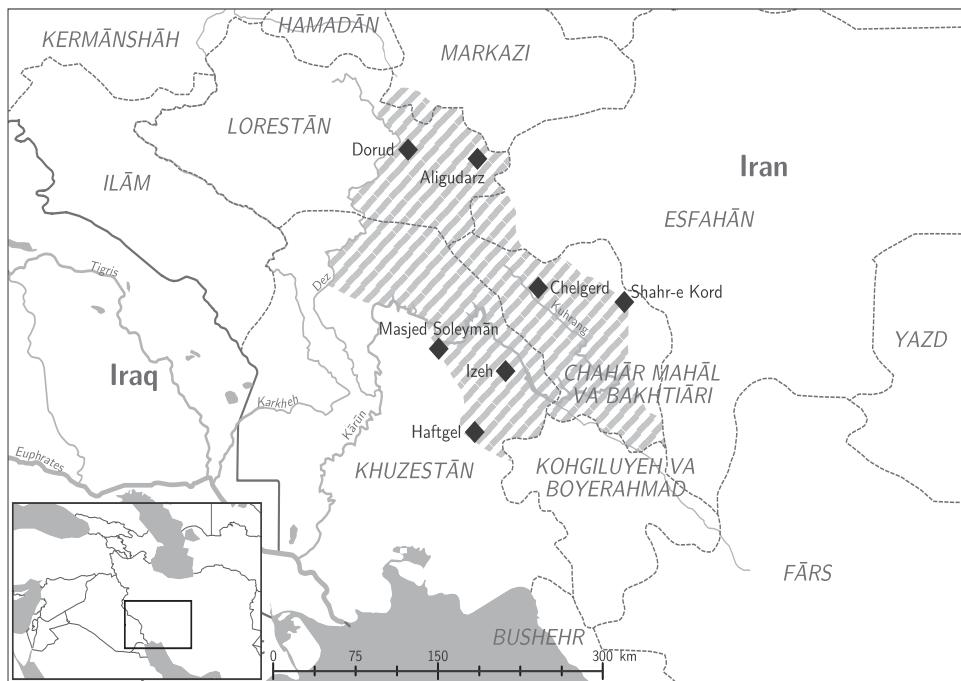


Figure 1: The Bakhtiari language area (based on AA 2014:17)

1919). While Kerimova's (1982) overview of Bakhtiari brings together and organizes a wide cross-section of data from these prior sources, Windfuhr's (1988) article concentrates on historical phonology and key morphological features of the language. Sociolinguistic background on the Bakhtiari language community and its place within the Lori language continuum are explored in Anonby (2003, 2012). A detailed phonological description, along with a long glossed text, verb paradigms, and a 1500-word lexicon, was published by Anonby and Asadi (AA 2014). Madadi (1996) and others have also produced lexicons of Bakhtiari, and Zolfaghari (2017) has written a multi-faceted study of the language, with a focus on lexicon, from an anthropological-linguistic perspective. An in-depth study of the geographic distribution of Bakhtiari in C&B Province has appeared in *ALI* (the online *Atlas of the Languages of Iran*, eds. Anonby and Taheri-Ardali 2015–18, hereafter AT 2015–18), and a dialectology of Bakhtiari dialects spoken there is underway in the same forum (Taheri-Ardali and Anonby 2017; Taheri-Ardali 2017). Schreiber et al. (2017) have looked at areal grammatical convergence between Bakhtiari and Turkic varieties in C&B. Finally, Khosravi (1996) has made extensive notes on the grammar of Bakhtiari, and E. Tāheri has provided the most complete grammatical description to date in his (2010) book.

Bakhtiari has a long and rich tradition of oral literature, including genres such as folktales, proverbs and sung poetry. There is no standardized orthography, but in the last century speakers have adapted Persian orthography to Bakhtiari and have published many volumes of oral literature (some well-known examples include: Dāvari 1964; Afsar Bakhtiyāri 1965; Forutan 1996; Ra’isi 1996; Kiyāni Haft Lang 2000; and Shirin Bayān 2007). Other recent publications in Bakhtiari include a collection of prose writings on language and culture (Be’vār 2000), a Bakhtiari adaptation of the *Shahnameh* epic (Khosraviniyā 2010) and a novel entitled *Brotherhood* (B. *gavgari*) (Khosraviniyā 2011).

As is the case for all languages in Iran, Persian – in its various historical incarnations and dialectal forms – has exerted a long-standing influence on Bakhtiari. Following many centuries where tendencies of divergence and convergence between Bakhtiari and its south-western sister Persian were more balanced, over the last century Bakhtiari structures at all levels of the language are increasingly hybridizing with their Persian counterparts under the influence of universal schooling and media: radio, television, and now internet. Perhaps even more significantly, parents in many parts of the language area, and especially urban locations, are teaching their children Persian as a first language in the home (Taheri-Ardali 2015; AT 2015–18). Consequently, despite a large language community, a pattern of language shift has been set in motion and the future of the language is uncertain.

The present study provides a linguistic overview of Bakhtiari, with an account of key elements of phonology §2, morphology §3, and syntax §4. This description is followed by a glossed text §5. The analysis here is based on the Masjed Soleymān dialect as described in Anonby and Asadi (AA 2014), unless otherwise noted, but salient points of dialectal variation in the language are addressed, and the glossed text is from the Ardal dialect of C&B Province.

2. Phonology

2.1. Segmental phonology

There is significant variation in the phonological inventories of dialects across the Bakhtiari language area. Urban varieties and Bakhtiari dialects in the east, which are transitional to the Persian varieties of the Iranian Plateau, tend to pattern more closely with Persian, whereas dialects in the west are more distinctive.

Some urban Bakhtiari varieties and dialects in the east exhibit a Persian-type vowel system with six vowels (Taheri-Ardali and Anonby 2017). In most dialects, however, the Middle Iranic *majhul* (long mid) vowels ē and (sometimes) ō (see Perry 1996) live on in various ways, whether as unitary mid-high vowels ē ō, diphthongs *ie ue*, or vowel-consonant sequences *ey ov* (contrasting with *ay av*)

(e.g., *zēne/zeyne/ziene* ‘woman’) (AA 2014: 56–58; Lorimer 1922: 23, 32). The historically long vowels *i*, *ā* and *u* – and, where they occur, *ē* and *ō* – still tend to be pronounced with greater phonetic length than the vowels *e*, *a* and *o*. This difference is not consistent in all contexts, though, and consequently vowel quality rather than length is the primary criterion for distinguishing among vowels (AA 2014: 60–61).

Table 1: Bakhtiari vowel inventory

<i>i</i> [i·]	<i>u</i> [u·]
(<i>ē</i> [e·]/etc.)	(<i>ō</i> [o·]/etc.)
<i>e</i> [ɛ]	<i>o</i> [ɔ̄]
<i>a</i> [a]	<i>ā</i> [v̄]

In the Bakhtiari dialect of Sar Āqā Seyyed in northern C&B Province, contrast between *u* and its fronted counterpart *ü* has also been observed, as evidenced by the minimal pair *du* ‘yogurt drink’ / *dü* ‘smoke’ (Taheri-Ardali 2017).

The Bakhtiari consonant inventory contains members from six general places of articulation, and six manners of articulation. It is more uniform across dialects than the vowel inventory, with the only significant case of variation being the absence of a voiceless uvular obstruent *q* in some eastern dialects; there, it is replaced by its voiced counterpart *g* (e.g., *qalb/galb* ‘heart’).

Table 2: Bakhtiari consonant inventory

	labial	alveolar	postalveolar	palatal/velar	uvular	glottal
vl. stop	<i>p</i>	<i>t</i>	<i>c</i> [tʃ]	<i>k</i> [k~k̄~c]	<i>q</i>	
vd. stop	<i>b</i>	<i>d</i>	<i>j</i> [dʒ]	<i>g</i> [g~ḡ~ɟ]	<i>g</i> [g~k̄]	
vl. fricative		<i>s</i>	<i>š</i> [ʃ]			<i>x</i> [χ]
vd. fricative		<i>z</i>	<i>ž</i> [ʒ]			
nasal	<i>m</i>	<i>n</i>				
lateral		<i>l</i>				
trill		<i>r</i> [r]				
glide	<i>v</i> [w~v~v̄]			<i>y</i> [j]		<i>h</i> [h~ħ]

Phonetically, vowel-initial words have a glottal stop [?] onset, but this segment is not contrastive ([?]av ‘water’, [?]eylevār ‘jaw’, [?]isā ‘you (pl.)’).

While *v* and *h* both have fricative realizations in some contexts, Anonby and Asadi (2014: 25) categorize them as glides because of their distribution and behaviour. The glide *v* in particular has three allophones [v], [v̄], [w] ([v]aro ‘high’, ko[v]ār ‘skeleton’, ša[w] ‘night’).

When *h* appears after vowels, it is realized phonetically as length on the preceding vowel: *peh* → *p[ɛ̄]*, *māh* → *m[ɔ̄]* ‘month’, *koh* → *k[ɔ̄]* ‘mountain’. The consonantal value of postvocalic *h* reappears weakly, as a voiced (i. e., breathy) [fi] when a vowel follows, and the vowel is shortened to its original duration (*ya kohey* → *ya k[ɔ̄fi]ey* ‘a mountain’).

Voiceless stops are non-contrastively aspirated in most positions ([p^h]ar ‘leaf’, *ma[k^h]al* ‘barren (land)’, *hayva[t^h]* ‘appearance’), but are not aspirated when they follow another obstruent (*poš[q]āv* ‘plate’) or precede another consonant (*qā[p]niden* ‘to grab’).

Velar stops *k* and *g* are palatalized [k^j], [g^j] – and in some dialects, fully palatal [c], [ɟ] – before front vowels ([k^j]ar/[c]ar ‘deaf’, [g^j]ap/[ɟ]ap ‘big’) and in codas (*hā[k^j]/hā[c]* ‘dirt’, *len[g^j]/len[ɟ]* ‘leg’).

The voiced alveolar stop *d*, which is pronounced as [d] word-initially and after a consonant ([d]ā ‘mother’, *gar[d]en* ‘neck’), is conspicuously realized as an inter-dental approximant [ð] after a word-internal vowel or glide, whether in the middle of a word (*ga[ð]e*, *meh[ð]e* – both ‘stomach’) and or word-finally (*ka[ð]* ‘waist’, *zey[ð]* ‘s/he hit’). In the Sar Āqā Seyyed dialect of north-western C&B, a complete phonemic shift to *r* has taken place in the same phonological position (*mehre* ‘stomach’, *zeyr* ‘s/he hit’; Taheri-Ardali 2017, AT 2017).

The voiced postalveolar fricative [ʒ] is contrastive in some dialects (B. of Masjed Soleymān *žerde* ‘torn’, *luže* ‘evading’, *gož* ‘anger’; AA 2014: 29–31, 46), but in many other dialects it only appears as an allophone of other postalveolar obstruents before another non-continuant segment (e. g., *hi[ʒ]dah* ‘eighteen’, *ga[ʒ]din* ‘scorpion’ *ma[ʒ]me/ma[ʒ]mah* ‘platter’; Taheri-Ardali 2017; AT 2017).

Geminated (i. e., doubled or strengthened) consonants are uncommon (*garr* ‘person with ringworm’, *takabbor* ‘arrogance’), and geminated consonants in Persian borrowings or cognates are usually simplified (P. *darre*, B. *dare* ‘valley’; P. *ammā*, B. *amā* ‘but’).

When *n* follows a vowel-glide sequence within a coda, it is realized as nasalization on this sequence (*zovn* → *z[ɔ̄w]* ‘tongue’, *dendown* → *dend[ɔ̄w]* ‘tongue’, *paynš=tā* → *p[ãj]š=tā* ‘five (count?’)). In eastern dialects where cognate sequences do not contain a glide (e. g., *zun* ‘tongue’, *dendun* ‘tooth’) *n* is retained as a full phonetic consonant.

2.2. Syllable and word structure

Syllables are composed of an optional simple onset, a nucleus, and an optional coda with up to three consonants. There are seven attested syllable shapes in Bakhtiari:

V	<i>i</i>	‘this’ (as in <i>i dast=o</i> ‘this hand’)
VC	<i>āš</i>	‘stew’
VCC	<i>asp</i>	‘horse’
CV	<i>jo</i>	‘barley’
CVC	<i>bur</i>	‘brown’
CVCC	<i>leng</i>	‘leg’
CVCCC	<i>be.peyžn</i>	‘twist, wrap (impv.)’

In inherited vocabulary, there are clear constraints on the voicing and sonority sequencing of coda consonants. In words borrowed from Arabic (via Persian), atypical coda clusters are in some cases retained (P.=B. *toxm* ‘egg’), but in other cases the clusters are reconfigured (P. *hokm* > B. *hoym* ‘order (n.)’, P. *fekr* > B. *ferg* ‘thought (n.)’). Codas may contain up to three consonants, but in this case the first consonant is always a glide *y v* or *h* (*šovnz.dah* ‘sixteen’, *be.seyvn* ‘stir! (sg.)’).

Vowel sequences are uncommon, with only *ai* (*baig* ‘bride’) and *āi* (*māi* ‘fish’) attested within a few inherited morphemes. Additionally, in some dialects, vowel sequences have emerged in borrowed words where the source language’s glottal stop, which is not phonemic in Bakhtiari, is lost (*sā?at* > B. *sāat* ‘hour’, P. *do?ā* > B. *doā* ‘spontaneous prayer’; but note that some other dialects ? is replaced by *h*: B. *sāhat*, *dohā*). At certain morpheme boundaries where vowels come together, vowel sequences are permitted (*be-* ‘(impv.)’ + *ašn* ‘hear’ → *beašn* ‘hear! (2sg.)’; *ya* ... = *ey* ‘(indef.)’ + *doro* ‘lie (n.)’ → *ya doroej* ‘a lie (n.)’). In most cases, however, the vowel sequence is eliminated through deletion of one of the vowels (*e-* ‘(impfv.)’ + *ašn* ‘hear’ + *-om* ‘1sg.’ → *ašnom* ‘I hear’; *pā* ‘foot’ + *=om* ‘1sg.’ → *pām* ‘my foot’) or the insertion of a glide between them (*taynā* ‘alone’ + *=om* ‘1sg.’ → *taynāyom* ‘I am alone’; *hālu* ‘paternal uncle’ + *=ay* ‘(voc.)’ → *hāluvay* ‘hey, paternal uncle!’).

There are no absolute restrictions on minimal lexical words, although words comprised of a single syllable ending in the historically “short” vowels *e*, *a* and *o* (see the vowel inventory in §2.1 above) are rare (AA 2014: 60–61); examples include *se* ‘three’ and *na* ‘no’ (see §3.2.3), and a few verb forms such as *go* ‘s/he said’ and *xa* ‘s/he ate’.

Monomorphemic words are almost always one or two syllables long, although semantically opaque three- and four-syllable nouns originating as compounds or borrowed words can be found (*ey.le.vār* ‘jaw’, *za.lā.la.ti* ‘distress’). In morphologically composite words, there can be more than four syllables (*i hov.ne.hā.ho.ne* ‘these houses (obj.)’, *na.ez.mā.yi.do.me.sovn* ‘I didn’t test them’).

2.3. Stress and intonation

In Bakhtiari, stress (or, more technically, phrasal accent) is associated with content words: verbs, nouns, adjectives and adverbs. It is normally hosted by the final syllable of the stem (*pā* ‘foot’, *bo tol* ‘beetle’, *kabu de* ‘poplar tree’). Affixes

and clitics attached to these words can be inherently stressed or unstressed (an inventory of affixes and clitics, along with their designation for stress, is given in Anonby and Asadi (AA 2014: 69–71)). Stressed morphemes override the underlying stress of the stem ('*pā* ‘foot’+='ke ‘(def.)’ → *pā'ke* ‘foot (def.)’; '*na-* ‘(neg.)’ + '*xa* ‘s/he ate’ → *naxa* ‘s/he didn’t eat’), whereas unstressed morphemes have no effect, even if they are word-final ('*pā* ‘foot’+='ne ‘(obj.)’ → *'pāne* ‘foot (obj.)’).

Intonation is marked in all sentence types through the raising of pitch on the stressed syllables of pragmatically salient words. Declarative sentences, commands, and content questions are marked by a general decline in pitch. In yes/no questions, the pitch is raised on the first pragmatically salient element of the sentence, and stays high until the end of the sentence, where it is raised even further. Further discussion of intonation in Bakhtiari, along with pitch traces for example sentences, is provided in Anonby and Asadi (AA 2014: 66–69).

3. Morphology

Nominals (§3.1) and verbs (§3.3) are open word classes in Bakhtiari. Closed classes (§3.2) include numerals, prepositions, pronouns, and a heterogeneous set of particles.

3.1. Nominals

Nominals, which include nouns (§3.1.1), adjectives (§3.1.2) and adverbs (§3.1.3), share many structural properties but can be distinguished through their combinatorial possibilities in morphology and their syntactic distribution.

3.1.1. Nouns

Nouns are inflected for number (plurality) and definiteness. Possession and other types of grammatical association are expressed through use of an *ezāfe* construction (§4.1) rather than case marking, but there is an object marker used with nouns (see in this section below). There is no grammatical gender.

Nominal plurality is marked with addition of the suffix *-(h)ā*, but its use is correlated to contexts where the plural noun is definite, or when plurality is being emphasized. The allomorph *-ā* is found with consonant-final stems (*čel-ā* ‘arms’, *merzeng-ā* ‘eyelashes’), and *-hā* is found with most vowel-final stems (*kolo-hā* ‘locusts’, *hovne-hā* ‘houses’). When the stem ends with a high vowel, however, the corresponding glide is inserted before *-ā* (*ti-y-ā* ‘eyes’, *soru-v-ā* ‘songs’).

The plural suffix *-ovn* is used in place of *-(h)ā* with a small set of nouns that includes handicapped persons (*kar-ovn* ‘deaf people’, *šal-ovn* ‘lame people’) as well as certain animals (*gusend-ovn* ‘sheep, goats’, *pāzen-ovn* ‘ibexes’).

A collective/plural suffix *-gal* (after alveolar consonants) / *-yal* (elsewhere) is used with a closed set of animate nouns (*kor-gal* ‘boys’, *boz-gal* ‘goats’, *ādom-yal* ‘people’, *hāšo-yal* ‘eagles’), as well as with nominalized adjectives referring to groups or people or animals (*tešne-yal* ‘the thirsty’, *zende-yal* ‘the living’). In southern dialects of C&B Province, which border on the Southern Lori language area, this suffix takes the form *-al* and is generalized to inanimate nouns as well (*deraxt-al* ‘trees’, *mel-al* ‘hairs’).

Definiteness and indefiniteness are both explicitly marked in the morphology. The definite suffix/clitic (which carries stress, but appears to be syntactically mobile; see AA 2014: 93) takes the form *=ke* with vowel-final nouns (*pā=ke* ‘the foot’, *hovne=ke* ‘the house’), and *=eke* with most nouns ending in two consonants (*hast=eke* ‘the bone’, *ahd=eke* ‘the treaty’). Stems ending in a single consonant can take either form, with the longer form being more emphatic (*sar=ke~sar=eke* ‘the head’, *tav=ke~tav=eke* ‘the fever’). This so-called ‘definiteness marker’ is not, however, simply the functional equivalent of the definite article in the languages of western Europe (for example). Investigation of Bakhtiari texts shows that this marker is often absent on noun phrases even when they are conceptually ‘definite’ (see, among others, Abbott 2004) and would require a definite article in other languages that mark definiteness. The precise nature of the definiteness marking in Bakhtiari is a topic for further research.

Indefiniteness is marked using the discontinuous morpheme *ya ...=ey* (*ya dovdar=ey* ‘a girl’, *ya pā=ey* ‘a foot’; *ya ...=i* in some dialects), with the initial *ya* component (cf. *yak* ‘one’) sometimes omitted. When this morpheme is attached to a stem ending in a high vowel, the corresponding glide is inserted (*ya geli-y=ey* ‘a throat’, *ya tu-v=ey* ‘a room’).

A demonstrative function is similarly marked on nouns using a discontinuous morpheme *i ...=(h)o* ‘this’ / *u ...=(h)o* ‘that’ (cf. the 3sg. distal/neutral free pronoun *ho*; see §3.2.4), with the allomorph *=o* used for consonant-final stems (*i/u dast=o* ‘this/that hand’, *i/u del=o* ‘this/that heart’) and *=ho* used for vowel-final stems (*i/u pā=ho* ‘this/that foot’, *i/u torne=ho* ‘this/that tuft of hair’).

Direct objects are usually marked with the clitic *=ne* (*ri=ne* ‘face (obj.)’, *say=ne* ‘dog (obj.)’), which is reduced to *=e* with consonants other than glides (*xayār=e* ‘cucumber (obj.)’, *gom=e* ‘footstep (obj.)’, *hars=e* ‘teardrop (obj.)’). When an object is also marked as indefinite, the indefinite enclitic precedes the object marker (*xayār=ey=ne* ‘a cucumber (obj.)’, and when marked as definite, the definite enclitic similarly precedes the object marker (*xayār=eke=ne* ‘the cucumber (obj.)’). However, in certain situations direct objects can be unmarked, e.g., *berenj* ‘rice’ in (58) below, and *gā* ‘cow’ in (84). The distribution and function of *=ne* are evidently distinct from *=rā*, as are the historical origins of the two object markers (the etymology of *=ne* is currently unclear; see the discussion in AA 2014: 89), but the contextual factors that govern the use of *=ne* have yet to be established.

Pronominal possession is marked by attaching an enclitic pronoun (see also §3.3.2 below) to the noun.

- (1) *kor=om* ‘my son’
boy=1SG
- (2) *ri=s* ‘his/her face’
face=3SG
- (3) *tu=tovn* ‘your (pl.) room’
room=2PL

Nominal possession is discussed along with noun phrase morphosyntax in §4.1 below.

3.1.2. Adjectives

There are many unitary adjectives in Bakhtiari, and these are structurally similar to nouns (*pati* ‘empty’, *taynā* ‘alone’, *zel* ‘curious’). Adjectives are productively derived from nouns using the suffix *-i* (*av* ‘water’, *av-i* ‘aquatic’; *taš* ‘fire’, *taš-i* ‘fiery, angry’). Nouns are also productively derived from adjectives using the same suffix form *-i* (*zende* ‘living’, *zende-i* ‘life’; *nā-xoš* ‘ill’, *nā-xoš-i* ‘illness’). Adjectives are joined to the noun they qualify using an *ezāfe* construction (see §4.1 below).

3.1.3. Adverbs

Adverbs are also structurally similar to nouns and adjectives. Whereas nouns and corresponding adjectives are distinct words that can be derived from one another, there is no formal difference between adjectives and corresponding adverbs (*xov* ‘good’, ‘well’; *tond* ‘fast’, ‘quickly’; *bahd* ‘subsequent’, ‘later’).

3.2. Closed word classes

3.2.1. Numbers

As in most languages, numbers (*yak* ‘one’, *do* ‘two’, *se* ‘three’, ...) form a clear closed morphological class in Bakhtiari. When a noun is explicitly counted, it is accompanied by a numeral classifier, most commonly *=tā* (*do=tā gavu* ‘two brothers’, *se=tā dadu* ‘three sisters’). There is no plural marking (cf. §3.1.1) on nouns accompanied by a number.

3.2.2. Prepositions

There is a large but closed set of prepositions in Bakhtiari. Many prepositions with a local sense are derived from body parts (*kel* ‘side’>‘beside’, *ri* ‘face’>‘on’, *pošt* ‘back’>‘behind’), and still function as nouns when used in the appropriate contexts. Some prepositions are composed of a simple preposition and another word (*men-jā* ‘between (lit. in-place)’, *vā-ri* ‘on top of (lit. with-face/on)’, *vā-pošt* ‘at the back of (lit. with-back/behind)’. These compound prepositions in particular are often used with an adverbial function as well:

- (4) *rah vā-pošt*
go.PST.3SG behind
's/he went backwards'

The *ezāfe* construction (§4.1) is used to join some prepositions to their complements:

- (5) *bahd=e ho* ‘after that’
after=EZ that
- (6) *men=e sok* ‘in the corner’
in=EZ corner

Pronominal complements of prepositions are marked using enclitic pronouns (§3.2.4).

- (7) *b=om/be=m* ‘to me’
to=1SG
- (8) *kel=et* ‘beside you (2sg.)’
beside=2SG
- (9) *si=s* ‘for him/her’
for=3SG
- (10) *ze=sovn* ‘from them’
from=3PL

3.2.3. Particles

There is a small set of particles with diverse functions. These items, which are comprised of single syllable with the historically short vowels *e*, *a* and *o*, can be pronounced as words on their own, for example as single-word answers. However, when they are used in a string of free speech they attach phonologically to the adjacent (usually following) word. Particles include the numbers *do* ‘two’ and *se* ‘three’, prepositions *be* ‘to’ and *ze* ‘from’, the first part of the discontinuous morphemes used for indefinites and demonstratives (§3.1.1), free pronouns (§3.2.4),

the interjection *na* (cf. §3.5), *pa* ‘then’, the coordinating conjunction (*v*)*o* (§4.3.1) and the subordinating conjunction *ke* ‘that’ (§4.3.2).

3.2.4. Pronouns and other person marking

Six person categories are marked in Bakhtiari morphology: first, second, and third person, each of which can be singular or plural. Third person free pronouns, both singular and plural, exhibit an obligatory proximal/distal distinction (thereby functioning as demonstrative pronouns at the same time; for the demonstrative function with nouns, see §3.1.1), although the distal form is used in neutral contexts as well. Table 3 below summarizes Bakhtiari person marking paradigms, and examples of each category are given in the relevant sections throughout this study.

Table 3: Person marking in Bakhtiari

	free pronoun	verb ending (perfective)	verb ending (non-past) and non-past indicative copula	verb ending (impera- tive)	enclitic pronoun (verbal object, pos- session, prepositional complement, etc.)
1SG	<i>mo</i>	-(<i>o</i>) <i>m</i>	-(<i>o</i>) <i>m</i>		=(<i>o</i>) <i>m</i>
2SG	<i>to</i>	- <i>i</i>	- <i>i</i>	-Ø	=(<i>e</i>) <i>t</i>
3SG	<i>ho</i> (dist.), <i>yo</i> (prox.)	-Ø	- <i>e</i>		=(<i>e</i>) <i>s</i>
1PL	<i>imā</i>	- <i>im</i>	- <i>im</i>	- <i>im</i>	=(<i>e</i>) <i>movn</i>
2PL	<i>isā</i>	- <i>in</i>	- <i>in</i>	- <i>in</i>	=(<i>e</i>) <i>tovn</i>
3PL	<i>ono</i> (dist.), <i>ino</i> (prox.)	-(<i>e</i>) <i>n</i>	-(<i>e</i>) <i>n</i>		=(<i>e</i>) <i>sovn</i>

In the table above, the vowels in parentheses are generally dropped when the morphemes accompany a vowel-final stem (see AA 2014: 69–90 for examples).

Reflexive pronouns are constructed using the root *xo* ‘self’ with enclitic pronouns (*xom* ‘myself’, *xosovn* ‘themselves').

3.3. Verbs

Bakhtiari verbs are of several lexical types. Some are simple (e. g., *avorden* ‘to bring’), but many also contain additional elements such as:

- a pre-verbal prefix *dar-/der-* (lit. ‘in’), *vā-* (‘with’), or *vor-* (‘on’): *der-avorden* ‘to remove’, *vā-biden* ‘to become’, *vor-čarden* ‘to crawl’;
- a pre-verbal nominal: *gir avorden* ‘to obtain’, cf. *gir* ‘hold (n.)’; *bālā avorden* ‘to vomit’, cf. *bālā* ‘above, up’; or

- a post-verbal prepositional phrase or adverb: *avorden belam* ‘to lower’, cf. *belam* ‘below, down’; *lağā zayden be gel* ‘to stomp’, cf. *lağah* ‘kick (n.)’, *zayden* ‘to hit’, *be* ‘to’, *gel* ‘mud’.

A number of verbs exhibit combinations of these structures (e. g., *dast vor-avorden* ‘to caress’, cf. *dast* ‘hand’; *šalāq kosten men=e guš* ‘to slap’, cf. *šalāq* ‘whip (n.)’, *kosten* ‘to pound’, *men=e* ‘inside of’, *guš* ‘ear’).

Each Bakhtiari verb has two stems: a non-past stem, which is used for non-past indicative, subjunctive, and imperative verb forms; and a past stem. Many past stems are the same or similar to their non-past counterparts, but with an additional *-id*, *-d* or *-est* ending (*davn-/dovnid-* ‘run (non-past/past)’, *vorčar-/vorčard-* ‘crawl (non-past/past)’, *dovn-/dovnest-* ‘know (non-past/past)’). For many other stems, the correspondence is partially irregular (*heył-/hešt-* ‘leave, let (non-past/past)’, *estovn-/estayd-* ‘get, obtain (non-past/past)’) and, in a few cases, suppletive (*-yā-/(o)veyd-* ‘come (non-past/past)’, *bin-/did-* ‘see (non-past/past)’). Several very common non-past stems comprise a single consonant (*-d-/dād-* ‘give (non-past/past)’, *-n-/nehād-* ‘put (non-past/past)’, *-r-/rahd-* ‘go (non-past/past)’; AA 2014: 125).

3.3.1. Non-finite verb forms

Non-finite forms include infinitives and perfect participles. Infinitives are formed by adding the suffix *-en* (*-an* in some eastern dialects) to the past stem (*xavsid-en* ‘to sleep’, *tāšid-en* ‘to shave’). Infinitives are used as verbal nouns, as the following example shows:

- (11) *gavu kočir-e ze zendei kerd-en sar tu ājez
brother young-DEF from life do-INF on room tired
ābid
become.PST.3SG*
‘The young[er] brother became tired of living on the roof.’

Perfect participles, formed by adding the suffix *-e* to the past stem (see the discussion of perfect verb forms under finite verb forms immediately below), are used as adjectives (*resid-e* ‘ripe’, cf. *resid-* ‘arrive, ripen (past)’; *gap ābid-e* ‘grown up’, cf. *gap ābid-* ‘grow, grow up (past)’).

3.3.2. Finite verb forms

Finite forms are inflected for person and TAM (tense/aspect/mood). There are many irregularities in verb conjugations (see AA 2014: 121–153), as well as many regular and irregular morphophonemic alternations (AA 2014: 69–90), and only the most important patterns are presented here.

Finite verb forms built on the non-past stem include the non-past (or “present”) indicative, subjunctive, and imperative.

The non-past indicative, which is used for basic present as well as habitual functions, varies significantly among Bakhtiari dialects. In the dialect of Masjed Soleyman (AA 2014), which is in focus here, this verb form is usually expressed with an obligatory imperfective prefix *e-* on the non-past stem, along with person-marking suffixes.

Table 4: Non-past indicative conjugation of ‘sleep’

1SG	<i>e-xavs-om</i>
2SG	<i>e-xavs-i</i>
3SG	<i>e-xavs-e</i>
1PL	<i>e-xavs-im</i>
2PL	<i>e-xavs-in</i>
3PL	<i>e-xavs-en</i>

For a few verbs in the Masjed Soleyman dialect, however, the prefix *e-* is absent or uncommon (e.g., *davn-om* ‘I know’, *tar-om* ‘I am able’, *torok-om* ‘I walk, wander’); and for another subset of verbs in the lexicon, the presence of this prefix appears to be correlated with an unmarked vs. continuous non-past distinction (e.g., *ašn-om/e-y-ašn-om* ‘I hear/I am hearing’, *davn-om~e-davn-om* ‘I run/I am running’, *heyl-om~e-heyl-om* ‘I leave (tr.)/I am leaving (tr.’), although the semantics of this distinction are neither clear nor consistent.

In some Bakhtiari dialects of C&B Province, the basic non-past indicative form is expressed using an unprefixed non-past stem, e.g., *xor-om* ‘I eat/I am eating’, *xor-i* ‘you (sg.) eat/you (sg.) are eating’, *xor-e* ‘s/he eats/s/he is eating’, etc. In the C&B dialect of Kuhrang, the prefix *e-* can also be used for the non-past indicative (Tāheri 2010: 163–164), although its possible semantic contribution – for example, to mark continuous or progressive aspect – is likewise unclear, as in Masjed Soleyman. In Ardal, C&B, where no cognate prefix is available for the non-past indicative, the unmarked vs. continuous/progressive distinction is signalled through the addition of the particle *hay*: *hay xor-om* ‘I am eating (cont./prog.)’ (Taheri-Ardali, field notes 2017).

As pointed out by van der Wal Anonby (pers. comm. 2017), the glossed text from Ardal in §5 below shows that the non-past can also be used with a mirative extension, that is, to mark “new, unexpected, and surprising information” (van der Wal Anonby 2011), somewhat analogous to the “dramatic present” in English (Huddleston & Pullum 2002: 129–131). Examples of this usage to mark an exciting episode in a personal account, repeated here (12–14) from the glossed text, occur in the middle of a string of past verb forms:

- (12) *hālā ke bin-e hay čerā takun xor-e!*
now COMPL see.NPST-3SG PROG stove shaking eat.NPST-3SG
‘now she sees that the stove is shaking!’
- (13) *tā virist-e dā en-e dov*
until stand.NPST-3SG mother put.NPST-3SG running
‘so mom stands up, she starts running’
- (14) *pa harči ogo b-iyav hune*
so whatever say.PST.3SG IMPV-come.IMPV.2SG house
polof-e!
collapse.NPST-3SG
she was saying over and over (lit. so she said whatever), “Come, the house
is collapsing!”

Among Iranic languages, this mirative use of the non-past differs from Persian and Tajik, where mirativity is an extended function of the perfect (Lazard 1985: 28–29, 2001: 361), and from Kumzari, where a dedicated mirative verb form has been identified (van der Wal Anonby 2011).

The subjunctive is marked, in Bakhtiari dialects generally, with a prefix *be-* and the same person-marking suffixes as are used for the non-past indicative (the data here is from AA 2014):

Table 5: Subjunctive conjugation of ‘sleep’

1SG	<i>be-xavs-om</i>
2SG	<i>be-xavs-i</i>
3SG	<i>be-xavs-e</i>
1PL	<i>be-xavs-im</i>
2PL	<i>be-xavs-in</i>
3PL	<i>be-xavs-en</i>

Some verbs are incompatible with the basic subjunctive conjugation and instead use a periphrastic perfect participle + subjunctive ‘be’ construction.

- (15) *tarest-e bu-m*
be_able-PF.PART be.SBJNC-1SG
‘that I be able...’

The imperative is formally equivalent to the subjunctive, except that its 2sg. form is marked with a null affix rather than the 2sg. suffix *-i*.

Table 6: Imperative conjugation of ‘sleep’

2SG	<i>be-xavs</i>
1PL	<i>be-xavs-im</i>
2PL	<i>be-xavs-in</i>

A few common verbs have irregular 2sg. imperative forms (*biyav/biyā* ‘come!’, *rav* ‘go!’).

Finite verb forms built on the past stem include the simple past, imperfective past, and the perfect. Person-marking suffixes on these forms are the same as those used with the non-past indicative (see above), except that the 3sg. suffix is a null morpheme.

For the simple past, the person-marking suffixes are simply attached to the past stem.

Table 7: Simple past conjugation of ‘sleep’

1SG	<i>xavsid-om</i>
2SG	<i>xavsid-i</i>
3SG	<i>xavsi(d)</i>
1PL	<i>xavsid-im</i>
2PL	<i>xavsid-in</i>
3PL	<i>xavsid-en</i>

When simple past stems end in a vowel+*d* or glide+*d* sequence, this final *d* is systematically omitted in the third person singular, depending on positional and discourse-related factors (discussed further in AA 2014:93–94, 127): *dā(d)* ‘s/he gave’, *di(d)* ‘s/he saw’, *rah(d)* ‘s/he went’, *zey(d)* ‘s/he hit’.

The imperfective past is formed by adding the prefix *e-*, which is also used for the non-past indicative (see above), to the past stem.

Table 8: Imperfective past conjugation of ‘sleep’

1SG	<i>e-xavsid-om</i>
2SG	<i>e-xavsid-i</i>
3SG	<i>e-xavsi(d)</i>
1PL	<i>e-xavsid-im</i>
2PL	<i>e-xavsid-in</i>
3PL	<i>e-xavsid-en</i>

The perfect is formed by adding the suffix *-e* after the person marking on the past stem (*rahd-om-e* ‘I have gone’, *did-im-e* ‘we have seen’). This contrasts with *ketābi* (literary) Persian, where the perfect suffix *-e* is attached directly to the stem (*ketābi P. raft-e-am* ‘I have gone’, *did-e-im* ‘we have seen’), but is matched by many varieties of Southern Kurdish, which have a similar ordering of morphemes, as well as in some NENA dialects, which likewise have a person ending followed by what is presumably a third person form of the copula (Geoffrey Khan, Geoffrey Haig, p.c.).

Table 9: Perfect conjugation of ‘sleep’

1SG	<i>xavsid-om-e</i>
2SG	<i>xavsid-i-y-e</i>
3SG	<i>xavsid-e</i>
1PL	<i>xavsid-im-e</i>
2PL	<i>xavsid-in-e</i>
3PL	<i>xavsid-en-e</i>

In some Bakhtiari varieties, the perfect is formed differently. In Ardal and some other areas of C&B Province, for example, it is marked in the third person singular with the suffix *-ak* (*resid-ak* ‘s/he has arrived, it has ripened’, cf. *resid* ‘s/he arrived, it ripened’), but for other persons there is no dedicated verb form and the simple past is used (*resid-om* ‘I have arrived’/‘I arrived’). Interestingly, in these dialects the perfect participle is formed with *-e* (not *-ak*), as is the case elsewhere in the language area (e. g., *resid-e* ‘ripe’; see further examples in the discussion of perfect participles in §3.3.1 above).

3.3.3. Complex verb forms

Important complex verb forms include the pluperfect and the progressive. In most dialects, the pluperfect is built by attaching the person-marking suffix directly to the past stem, followed by an invariable auxiliary *bi* (cf. *bi* ‘he/she/it was’). Parallel to what happens with the perfect, this structure differs from other Iranic languages such as Persian: there, person is marked on the ‘be’ auxiliary rather than the stem. The patterning of the Bakhtiari pluperfect is illustrated in the following example by the form *jam ābid-im bi* ‘we had gathered’:

- (16) *gav-yal dafe niyā=i*
brother-PL time previously=ADJZR
ke davr=e yak jam ābid-im bi
COMPL around=EZ one gathered become.PST-1PL be.AUX
manesti ādomizād ičo bi o
as.if human here be.PST.3SG and
harf-ā=movn=e ašni o
word-PL=1SG=OBJ hear.PST.3SG and
‘Brothers! The previous time, when we had gathered together, it seems a human was here and heard our conversation...’ (AA 2014: 94–95)

The formation of the progressive varies greatly among Bakhtiari varieties, with various combinations of auxiliary, verbal repetition, and a dedicated progressive adverbs. Progressive constructions in western dialects of Bakhtiari have not been studied, but in varieties of C&B Province in the east, some examples of the past progressive construction in particular, used for the construction ‘I was going’, are as follows (Schreiber et al. 2017; Taheri-Ardali, field notes 2016):

Table 10: Attested past progressive constructions in several dialects of C&B Province:
‘I was going’

<i>rahd-om</i>	<i>be-r-om</i>	(Ardal)
go.PST-1SG	SBJNC-go.NPST-1SG	
<i>hey</i>	<i>rahd-om</i>	(Ardal)
PROG	go.PST-1SG	SBJNC-go.NPST-1SG
<i>hey</i>	<i>rah-m</i>	(Lordegān)
PROG	go.PST-1SG	SBJNC-go.NPST-1SG
<i>i-rahd-om</i>	<i>be-r-om</i>	(Ardal, Lordegān)
IMPFV-go.PST-1SG	SBJNC-go.NPST-1SG	
<i>i-rah-m</i>	<i>be-r-om</i>	(Dastenā)
IMPFV-go.PST-1SG	SBJNC-go.NPST-1SG	
<i>dāšt-om</i>	<i>i-rah-m</i>	(Dastenā)
have.PST-1SG	IMPFV-go.PST-1SG	
<i>dāšt-om</i>	<i>i-rah-m</i>	(Dastenā)
have.PST-1SG	IMPFV-go.PST-1SG	SBJNC-go.NPST-1SG

3.4. The copula

In its non-past indicative form, the copula is expressed with the same person-marking morphemes as those used for non-past indicative verb forms.

Table 11: Non-past indicative copula with *zarākāl* ‘farmer’

1SG	<i>zarākāl=om</i>	‘I am a farmer’
2SG	<i>zarākāl=i</i>	‘you (sg.) are a farmer’
3SG	<i>zarākāl=e</i>	‘s/he is a farmer’
1PL	<i>zarākāl=im</i>	‘we are farmers’
2PL	<i>zarākāl=in</i>	‘you (pl.) are farmers’
3PL	<i>zarākāl=en</i>	‘they are farmers’

When the non-past indicative copula is used with a free pronoun (including demonstrative pronouns) or an interrogative pronoun, *n* is inserted between the two parts: *mo-n=om* ‘it’s me’, *yo-n=e* ‘it’s him/her/it/this one’, *isā-n=in* ‘it’s you (pl.)’, *hami=ho-n=e* ‘it’s that very one’, *koye-n=im* ‘where are we?’, *key-n=in?* ‘who are you (pl.)’, *kay-n=e* ‘when is it?’. Alternative explanations for this structure, including a possible historical relation to the differential object marker =*ne* (see §3.1.1 above), are discussed in AA 2014: 88–89.

For other verb forms (subjunctive, imperative and past), the verb stem *bu-/bid-* (‘be’, non-past/past) is used (*zarākāl bid-i* ‘you (sg.) were a farmer’, *zarākāl bu-y* ‘that you (sg.), be a farmer’, *zarākāl bu* ‘(you sg.) be a farmer!’). A full conjugation of the verbal copula, which shows significant variation depending on the speaker as well as the discourse context, is as follows:

Table 12: Conjugation of the verbal copula *bu-/bid-* ‘be’

	subjunctive	imperative	past
1SG	<i>bu-m/bu-om/bu-v-om</i>		<i>bid-om</i>
2SG	<i>bu-y/bu-i/bu-v-i</i>	<i>bu</i>	<i>bid-i</i>
3SG	<i>bu/bu-e/bu-v-e</i>		<i>bi/bid</i>
1PL	<i>bu-ym/bu-im/bu-v-im</i>	<i>bu-ym/bu-im/bu-v-im</i>	<i>bid-im</i>
2PL	<i>bu-yn/bu-in/bu-v-in</i>	<i>bu-yn/bu-in/bu-v-in</i>	<i>bid-in</i>
3PL	<i>bu-n/bu-en/bu-v-en</i>		<i>bid-en</i>

3.5. Verbal negation

For all verb forms, verbal negation is marked on the inflected stem with the negative prefix *na-*. Its patterning is straightforward in non-prefixed indicative forms, such as the simple past, where *na-* is simply attached directly to the stem:

Table 13: Basic verbal negation

verb		negative form	
<i>neyašt-om</i>	'I looked'	<i>na-neyašt-om</i>	'I didn't look'
<i>torokest-om</i>	'I walked'	<i>na-torokest-om</i>	'I didn't walk'
<i>xavsid-om</i>	'I slept'	<i>na-xavsid-om</i>	'I didn't sleep'

There are a few common verbs whose initial stem consonant is deleted, however, sometimes along with other irregular alternations, when *na-* is attached.

Table 14: Contracted negative verb forms

<i>dovn-om</i>	'I knew'	<i>na-vn-om</i>	'I didn't know'
<i>did-om</i>	'I saw'	<i>na-yd-om</i>	'I didn't see'
<i>hed-om</i>	'I am (there)'	<i>n-eyd-om</i>	'I am not (there)'

With subjunctive as well as imperative verb forms, *na-* displaces the subjunctive/imperative prefix *be-* (see above).

Table 15: Negative subjunctive verb forms

<i>be-neyar-e</i>	'that s/he look'	<i>na-neyar-e</i>	'that s/he not look'
<i>be-torok-e</i>	'that s/he walk'	<i>na-torok-e</i>	'that s/he not walk'
<i>be-xavs-e</i>	'that s/he sleep'	<i>na-xavs-e</i>	'that s/he not sleep'

Table 16: Negative imperative verb forms

<i>be-neyar</i>	'look! (sg.)'	<i>na-neyar/mah-neyar</i>	'don't look! (sg.)'
<i>be-torok</i>	'walk! (sg.)'	<i>na-torok/mah-torok</i>	'don't walk! (sg.)'
<i>be-xavs</i>	'sleep! (sg.)'	<i>na-xavs/mah-xavs</i>	'don't sleep! (sg.)'

As shown in Table (16) above, the prefix *mah-* is used alongside *na-* in prohibitive (negative imperative) constructions.

When *na-* is attached to stems that begin with the non-past indicative prefix *e-*, the two morphemes fuse and are realized as *ney-*.

Table 17: Negative non-past indicative verb forms

<i>e-xavs-om</i>	'I sleep'	<i>ney-xavs-om</i>	'I don't sleep'
<i>e-d-om</i>	'I give'	<i>ney-d-om</i>	'I don't give'
<i>e-ger-om</i>	'I take'	<i>ney-ger-om</i>	'I don't take'

The negative prefix is in some cases found before a preverb (*na-dar-garnid* ‘s/he didn’t turn over (tr.)’, *na-der-avord* ‘s/he didn’t bring’), and in other cases it intervenes between the preverb and the stem (*dar-na-borid* ‘s/he didn’t protest’, *der-na-xard* ‘s/he didn’t turn around’). The placement of the prefix is therefore lexically determined, and depends on the internal morphological structure of the verb: it may even pattern in both ways when a single stem has two distinct meanings (*na-da-rahd* ‘it didn’t set (sun)’, *dar-na-rahd* ‘s/he didn’t turn around’; AA forthcoming).

3.6. Pronominal verbal objects

Pronominal verbal objects are normally marked with enclitic pronouns, which are hosted by the inflected verb. Table 18 gives the full paradigm of object clitics, following a third-person singular simple past verb. Table 19 illustrates clitic objects in combination with different values for subject marking. In light verb constructions, the clitic pronoun is hosted by the nominal element (see (20) below):

Table 18: Pronominal object marking on 3sg. past verb forms

<i>did=om</i>	‘s/he saw me’
<i>did=et</i>	‘s/he saw you (sg.)’
<i>did=es</i>	‘s/he saw him/her/it’
<i>did=emovn</i>	‘s/he saw us’
<i>did=etovn</i>	‘s/he saw you (pl.)’
<i>did=esovn</i>	‘s/he saw them’

Table 19: Pronominal object marking with different person values for subject

<i>did-om=et</i>	‘I saw you (sg.)’
<i>did-i=m</i>	‘you (sg.) saw me’
<i>did-im=es</i>	‘we saw him/her/it’
<i>did-im=etovn</i>	‘we saw you (pl.)’
<i>did-in=esovn</i>	‘you (pl.) saw them’
<i>did-en=emovn</i>	‘they saw us’

We are not aware of any significant differences between the tenses with regard to the deployment of object clitic pronouns; that is, the paradigms for clitic pronouns given in Tables 18 and 19 are essentially the same with a non-past verb form as host. Bakhtiari thus seems to lack tense sensitivity in alignment, resembling Persian in this regard (Haig 2008), though more research on Bakhtiari syntax is needed.

When both direct object and indirect object are found as pronouns, there is a constraint against marking both with enclitic pronouns: typically, one of the

objects is dropped (17, 18) or the direct object is retained as a full pronoun in pre-verbal position (19).

- (17) *dād-om=es*
give.PST-1SG=3SG
'I gave [the recipient] it' or 'I gave him/her [it]'
- (18) *dād-om=et*
give.PST-1SG=2SG
'I gave you [it]'
- (19) *ho=ne dād-om=et*
3SG.DIST=OBJ give.PST-1SG=2SG
'I gave you it'

In light verb constructions, the enclitic pronouns are attached to the nominal component of the construction.

- (20) *ravne=s kerd*
sending_away=3SG do.PST.3SG
's/he sent him/her away' (AA 2014: 99)

4. Syntax

The present study focuses on three topics central to Bakhtiari syntax: the *ezāfe* construction (§4.1), clause constituent order (§4.2), and clause linkage (§4.3), which includes coordination and subordination. Other constructions with importance for both syntax and morphology have been integrated into the preceding discussion; for example, noun affixes marking definiteness, demonstrative functions, and differential objects are discussed under noun morphology (§3.1.1).

4.1. The *ezāfe* construction

The *ezāfe* (associative) construction is used in Bakhtiari to form complex phrases of various types, including noun+noun, noun+adjective, and preposition+adjective phrases. The phrasal head is always on the left periphery of the phrase. When the noun phrase head is consonant-final, the associative morpheme =e is attached to it.

- (21) *dovdar=e pādešā* 'king's daughter'
daughter=EZ king

When the noun phrase head is vowel-final, however, there is no overt associative morpheme.

- (22) *lovne peleštuk* ‘swallow (bird) nest’
nest swallow

Noun+adjective (23, 24) as well as preposition+noun (25, 26) *ezāfe* constructions are formed in the same way.

- (23) *xayāl=e āside* ‘peace of mind (lit. unworried thoughts)’
thought=EZ unworried
- (24) *gā gap* ‘big cow’
cow big
- (25) *neheng=e aftavdarov* ‘close to sunset’
near=EZ sunset
- (26) *ze zemin* ‘from [the] ground’
from ground

When the non-head element in such constructions is a pronoun rather than a noun, enclitic pronouns, which attach directly to the phrasal head, are used (§3.2.2, §3.2.4).

4.2. Clause constituent order

For clauses with an overt subject (whether a noun or a free pronoun) and a verb, the default order is SV (subject-verb).

- (27) *piyā rāhd*
man go.PST.3SG
‘[the] man went’
- (28) *mo rāhd-om*
I go.PST-1SG
‘I went’ (with “I” as a new topic)

When a noun or pronoun has already appeared as the subject of a previous sentence, it is typically dropped in the subsequent sentence. The identity of the subject is still clear, however, from person-marking on the verb.

- (29) *rāhd*
go.PST.3SG
‘s/he went’
- (30) *rāhd-om*
go.PST-1SG
‘I went’ (with subsequent mention of “I”)

In clauses with a nominal object, the default order is (S)OV (31–33); the direct object marker *=ne* is introduced in §3.1.1 above.

- (31) *piyā hovne=ne xeri*
man house=OBJ buy.PST.3SG
'[the] man bought the house' (cf. AA 2014: 113)
- (32) *dovar=eke sive=ka=ne xa*
girl=DEF apple=DEF=OBJ eat.PST.3SG
'the girl ate the apple' (Ardal dialect)
- (33) *hovne=ne xerid-om*
house=OBJ buy.PST-1SG
'I bought [the] house'

Nominal subjects and objects that are emphasized, however – for example, in contrast with another possible noun phrase – may move to post-predicate (i. e., post-verbal) position.

- (34) *sive=ka=ne xa dovar=eke*
apple=DEF=OBJ eat.PST.3SG girl=DEF
'the girl ate the apple' (Ardal dialect)
- (35) *dovar=eke xa sive=ka=ne*
girl=DEF eat.PST.3SG apple=DEF=OBJ
'the girl ate the apple' (Ardal dialect)
- (36) *xa sive=ka=ne dovar=eke*
eat.PST.3SG apple=DEF=OBJ girl=DEF
'the girl ate the apple' (Ardal dialect)

Pronominal objects, which are expressed using enclitic pronouns attached to the predicate, are discussed separately in §3.6.

The indirect object of a verb of transfer usually follows the predicate. The word *das* (lit., 'hand') optionally introduces the recipient.

- (37) *setāre čağu=ne dā (das) Borjali*
Setare knife=OBJ give.PST.3SG (hand) Borjali
'Setare gave Borjali the knife' (Ardal dialect)

As is the case for recipients, goals of motion verbs likewise follow the predicate, as in the following (see (43) for additional post-predicate goals).

- (38) *rahd sar=e tu*
go.PST.3SG head/on=EZ room/house
's/he went onto the roof (lit. head of [the] room/on [the] room)'
(AA 2014: 107)

Addressees, however, precede the verb and are introduced with the preposition *ze* ‘from’ (*ez* in some dialects). (Further examples of citations are found in the discussion on subordination in §4.3.2 below.).

- (39) *setāre ez Borjali ogo unun duš ičo bi-y-en*
 Setare from Borjali say.PST.3SG 3PL.DIST yesterday here
be.PST-y-3PL

‘Setare said to Borjali, “They came here yesterday”.’ (Ardal dialect)

With regard to post-predicate elements, Bakhtiari thus follows the pattern of Vafsi (Stilo 2010) and the northwesterly dialects of Northern Kurdish (Haig 2017), in that recipients and goals of verbs of motion follow the predicate, while addressees precede the predicate.

4.3. Linkage

4.3.1. Coordination

A range of constructions are coordinated with the particle (*v*)*o* ‘and’. This particle (cf. §3.2.3 above), which can attach phonologically to the preceding or the following word, is transcribed here separately. When (*v*)*o* attaches to a vowel-final word, it is realized as *vo* (*dā vo bavu* ‘parents (lit. mother and father)'), and it appears as *o* elsewhere.

In the following example (40), (*v*)*o* is used to coordinate noun phrases (here represented by the nouns *bamr o pelang* ‘tiger and leopard’), verb phrases (*gā gapeyne šekāl kerden o avorden* ‘hunted and brought a big cow’), and clauses (*bamr o pelang rahden o gā gapeyne šekāl kerden o avorden* ‘tiger and leopard went and hunted and brought a big cow’).

- (40) *bamr o pelang rahd-en o gā gap=ey=ne*
 tiger and leopard go.PST-3PL and cow big=INDEF=OBJ
šekāl kerd-en o avord-en
 hunting do.PST-3PL and bring.PST-3PL
 ‘[The] tiger and [the] leopard went and hunted a brought [back] a big cow.’
 (AA 2014: 101)

The particle (*v*)*o* can also be used to coordinate independent sentences, as the following example (41) shows:

- (41) *hakāat=e xers ke tamom ābi*
 story=EZ bear COMPL finished become.PST.3SG
šeyr go
 lion say.PST.3SG
hame=movn ājez=im.
 all=1PL tired=be.1PL
 ‘When [the] bear’s story came to end, the lion said: “We are all tired”.’
si amšav diye bas=e
 for tonight other enough=be.3SG
 ‘“That’s enough for tonight”.’
o hame voreystā-n xavsid-en.
 and all get_up.PST-3PL sleep.PST-3PL
 ‘And they all got up and went to sleep.’

The end of the last sentence of (41) (*voreystān xavsiden* ‘they got up, they went to sleep’) also shows that verb phrases or clauses, which are in this case indistinguishable, can be coordinated through juxtaposition.

The conjunction *amā* ‘but’ is used for adversative coordination of clauses (42) as well as sentences (43).

- (42) *yak=i ze i do gavu*
 one=INDEF from this two brother
ādom=e sāf o sādoq=ey bi,
 person=EZ smooth and honest=INDEF be.PST.3SG
amā u yak=i rendal bi.
 but that one=INDEF crafty be.PST.3SG
 ‘One of these brothers was a simple person, but the other one was crafty.’
 (AA 2014: 96)

- (43) *aval=es gemovn kerd*
 first=3SG imagination do.PST.3SG
doz-ovn ovayd-en-e men=e āsiyav
 robber-PL come.PST-3PL-PRF in=EZ mill
tā čiy-ā dozi=ne bayn=e xo=sovñ
 until thing-PL stolen(adj.)=OBJ between=EZ self=3PL
bahr bo-kon-en.
 portion SBJNC-do.NPST-3PL
 ‘First, he thought the robbers had come inside the mill to divide the stolen things among themselves.’

- amā xov ke niyašt,*
 but well COMPL look.PST.3SG
di ya ruvā=yey ovayd vāibā
 see.PST.3SG a fox=INDEF come.PST.3SG within
o vā din=es āsiyav=ne jāru kerd o
 and with tail=3SG mill=OBJ broom do.PST.3SG and
nešast ya guše=ey.
 sit_down.PST.3SG a corner=INDEF
- ‘But when he looked carefully, he saw a fox that came inside and swept the mill with its tail and sat down in a corner.’ (AA 2014: 100–101)

4.3.2. Adverbial clauses

In narrative texts, clauses that depict a prior event or condition commonly precede the main clause. They often contain the particle *ke*, which then immediately precedes the verb of the adverbial clause. This may be seen in the preceding example (43) (*amā xov ke niyašt* ‘but when he looked closely’), as well as the following sentence (44):

- (44) *ya modat=ey ke godašt*
 a while=INDEF COMPL pass.PST.3SG
gavu kočir=e [...] ājez ābid
 brother young=DEF [...] tired become.PST.3SG
 ‘a while having passed, the younger brother [...] became tired’
 (AA 2014: 98–99)

Other adverbial complementizers include simple conjunctions such as *ayar* ‘if’ (45) or more complex conjunctive constructions – often containing *ke*: *vaxtey ke* ‘when (lit. a time that)’ (46) or *bahde yo ke* ‘after (lit. after this that)’ (47), which occur at the beginning of the adverbial clause.

- (45) *ayar ādomizād dovnest*
 if human know.PST.3SG
e-rahd o u hovne=ne e-xeri o
 IMPFV-go.PST.3SG and that house=OBJ IMPFV-buy.PST.3SG and
 ‘If a human knew, s/he would go and would buy that house ...’ (AA
 2014: 105–106)

- (46) *vaxt=ey ke hame xov seyr ābid-en,*
 time=INDEF COMPL all good full become.PST-3PL
seyr go
 lion say.PST.3SG
xāv amšav key hakāat=e aval=e e-go?
 well tonight who story=EZ first=OBJ IMPFV-say.NPST.3SG
 ‘When all of them became completely satisfied, the lion said: “Well, who is going to tell the first story tonight?”’ (AA 2014: 101–102)
- (47) *bahd=e yo ke hakāat=e gorg be āxer rasi,*
 after=EZ 3SG.PROX COMPL story=EZ wolf to end arrive.PST.3SG
xers sine xo=ne sāf kerd
 bear chest self=OBJ smooth do.PST.3SG
 ‘After the wolf’s story came to end, the bear cleared its throat...’ (AA 2014: 105)

While many adverbial clauses precede the matrix clause, as shown in the preceding examples, clauses expressing purpose or subsequent action are found after the matrix clause.

- (48) *piyā rah*
 man go.PST.3SG
tā rasi be jā=yey
 until arrive.PST.3SG to place=INDEF
ke xers god-e bi.
 COMPL bear say.PST.3SG-PRF be.AUX
 ‘The man went until he got to the place that the bear had said.’

Direct speech (49, 50) and perceptual complement clauses (51) also follow the matrix clause. Both are introduced without a complementizer.

- (49) *pah go biy-av*
 so say.PST.3SG IMPV-come.IMPV.2SG
 ‘So he said, “Come...”’ (AA 2014: 96–97)
- (50) *ho ogo amšaw iyā-m*
 3SG.DIST say.PST.3SG tonight come.NPST-1SG
 ‘S/he said, “I am coming tonight”.’ (Ardal dialect)
- (51) *di ze ro tu tā āsemovn gaf-tar=e.*
 see.PST.3SG from top room until sky big-CMPV=be.3SG
 ‘He saw that [the part] from the roof to the sky was larger’. (AA 2014: 98)

Semi-direct (52) and indirect (53) citations, on the other hand, use the particle *ke*.

- (52) *ho go ke amšaw iyā-m*
 3SG.DIST say.PST.3SG that tonight come.NPST-1SG
 ‘S/he said that s/he is coming tonight.’ (semi-direct speech) (Ardal dialect)
- (53) *ho go ke amšaw iyā*
 3SG.DIST say.PST.3SG that tonight come.NPST.3SG
 ‘S/he said that s/he is coming tonight.’ (indirect speech) (Ardal dialect)

5. Glossed text: zelzeleā ‘The earthquake’

The following text is from the dialect of Ardal, in C&B Province. Ardal is at the eastern end of the Bakhtiari language area and, as shown in §2.1 and elsewhere above, the dialects of this region display some additional features in common with Persian and other Southwestern dialects of the Iranian Central Plateau. The text is a personal account told by Homa Asadi-Ardali, a 60 year-old woman, who describes her experience during an earthquake forty years earlier, in 1977. Here, she is speaking to her nephew in the context of a larger audience of relatives.

- (54) *hebdah ruz bahd ez ayd bi bav=m,*
 seventeen day after from festival be.PST.3SG father=1SG
sāat čār bi.
 hour four be.PST.3SG
 ‘It was seventeen days after [the Now Ruz] festival, my dear (lit. my father), it was four o’clock.
- (55) *čār bi, di šom=emun=am nād-im –*
 four be.PST.3SG again dinner=1PL=ADD put.PST-1PL
 It was four, and we had put dinner on [to cook] –
- (56) *uso dovre čerā bid=o gāz.kapsuli bi –*
 then period stove be.PST.3SG=and gas.cylinder be.PST.3SG
 back then it was the time of stoves (i. e., burners) and gas cylinders –
- (57) *nād-im=eš sar čerā.*
 put.PST-1PL=3SG on stove
 we put it on the stove.
- (58) *gušt=en=am sohr kerd māmā kemutar,*
 meat=OBJ=ADD red make.PST.3SG grandmother Kemutar
xāst berenj bo-kon-e pā=š.
 want.PST.3SG rice SBJNC-make.NPST-3SG foot=3SG
 Grandma Kemutar fried the meat and was going to make rice to go with it.

- (59) *di ya takun=i* ...
 again INDEF shaking=INDEF
 Then a shaking ...
- (60) *hālā ke bin-e hay čerā takun xor-e!*
 now COMPL see.NPST-3SG PROG stove shaking eat.NPST-3SG
 now she sees that the stove is shaking!
- (61) *ogod-om dā, hay čerā takun xor-e!*
 say.PST-1SG mother PROG stove shaking eat.NPST-3SG
 I said, “Mom, the stove is shaking!”,
- (62) *ogo na! zelzele-ā gedim=e,*
 say.PST.3SG no earthquake-PL ancient=3SG
 she said, “No! Earthquakes are [something that used to happen] long ago”,
- (63) *hay ogo barkat.*
 PROG say.PST.3SG blessing
 she was saying, “Bless me (lit. blessing)”.
- (64) *ogod-om hay dā barkat di xelās –*
 say.PST-1SG PROG mother blessing again finished
 I said, “Mom, right now that blessing belongs to the past (lit. blessing finished again) –
- (65) *yo hay takun ed-e!*
 3.SG.PROX PROG shaking give.NPST-3SG
 it's shaking!”
- (66) *tā virist-e dā en-e dov*
 until stand.NPST-3SG mother put.NPST-3SG running
er-e duven,
 go.NPST-3SG down/outside
 So mom stands up, she starts running, she goes down outside,
- (67) *ogod-om dā!,*
 say.PST-1SG mother
 I said, “Mom!”,
- (68) *pa harči ogo b-iyav hune*
 so whatever say.PST.3SG IMPV-come.IMPV.2SG house
polof-e!
 collapse.NPST-3SG
 she was saying over and over (lit. so she said whatever), “Come, the house is collapsing!”

- (69) *ogod-om na! tā mahtāb=e na-yār-om,*
 say.PST-1SG no until Mahtab=OBJ NEG-bring.NPST-1SG
n-iyā-m.
 NEG-come.NPST-1SG
 I said, “No! If I don’t take Mahtab out, I won’t come.
- (70) *mahtāb hamučo vas mi ārd-ā.*
 Mahtab right_there fall.PST.3SG in flour-PL
 Mahtab fell right there into the flour.
- (71) *tāpu ārd-i=mun katāyun mi ya ini*
 clay.pot flour-ADJZR=1PL Katayun in one what’s_it_called
bi jā ... komod-i,
 be.PST.3SG place dresser-INDEF
 Our clay pot for flour, Katayun [an addressee in the audience], was in a
 what’s-it-called, clothes ... alcove (lit. place ... of dresser),
- (72) *ovay vas pišgavi.*
 come.PST.3SG fall.PST.3SG spread_all_over.PST.3SG
 it happened that it had fallen and spilled out all over.
- (73) *mahtāb fagir vast min=eš,*
 Mahtab poor fall.PST.3SG in=3SG
 Mahtab, poor thing, had fallen into it,
- (74) *mahtāb=e bord-om,*
 Mahtab=OBJ carry.PST-1SG
 I took Mahtab,
- (75) *gand-ā=ne bord-om,*
 rock.sugar-PL=OBJ carry.PST-1SG
 I took rock sugar chunks,
- (76) *legan.nuni=ne bord-om ...*
 bread.box=OBJ carry.PST-1SG
 I took the bread box ...
- (77) *šav ogo vā na-vāst-in.*
 night say.PST.3SG must NEG-stand.NPST-2PL
 [When it was] night, they (lit. s/he) said, “You (pl.) mustn’t stay”.
- (78) *hami inā ke bid-en*
 this_same these COMPL be.PST-3PL
ovayd-en si masan komak bo-kon-en,
 come.PST-3PL for for_example help SBJNC-do.NPST-3PL
 All these [people] who were there came to, you know, help,

- (79) *go na-vāst-in ičo vā be-r-in*
 say.PST.3SG NEG-stand.NPST-2PL here must IMPV-go.NPST-2PL
sar šāzeaziz.
 head/on Shaze.Aziz
 they said, “You mustn’t stay here, you have to go over to Shaze-Aziz [shrine]”.
- (80) *šav rahd-im sar šāzeaziz anul faġir=am bi.*
 night go.PST-1PL head/on Shaze.Aziz Anul poor=ADD be.PST.3SG
 In the night we went over to Shaze-Aziz, poor Anul was [with us] as well.
- (81) *ez hamī šomi hay zemin takun xa*
 from this_same dinnertime PROG earth shaking eat.PST.3SG
xodā=am hay bārun zay.
 God=ADD PROG rain hit.PST.3SG
 Right from dinnertime the earth was shaking, rain was pouring down (lit. God was raining).
- (82) *hay zayd zayd zayd tā dam=e*
 PROG hit.PST.3SG hit.PST.3SG hit.PST.3SG until beginning=ez
sohv ...
 morning
 It was raining, raining, raining till the break of dawn ...
- (83) *gusiyal=mun=am mi madrese.kār.dāneš i.duven bid-en.*
 sheep.PL=1PL=ADD in technical.school down be.PST-3PL
 The sheep were in the technical school down below.
- (84) *dah-bis=tā miš dāšt-im do-se=tā gā*
 ten-twenty=NUM ewe have.PST-1PL two-three=NUM cow
dāšt-im, do ... ya jof varzā dāšt-im ...
 have.PST-1PL two one pair bulls have.PST-1PL
 We had ten or twenty ewes, we had two or three cows, two ... we had one pair of oxen ...
- (85) *bāvā hamziy=am šixun bi.*
 grandfather Hamzi=ADD Shikhun be.PST.3SG
 Grandpa Hamzi was in Shikhun.
- (86) *šixun bi hā hā ... be.bahd ...*
 Shikhun be.PST.3SG yeah yeah afterward
 He was in Shikhun, mhmm ... after ...

- (87) *gusiyal=i ke vel kerd-im xo=šun rahd-en.*
sheep.PL= SPEC COMPL release do.PST-1PL self=3PL go.PST-3PL
When we freed the sheep, they themselves went out.
- (88) *ra[hd-en] harki bord=ešun min amu*
go.PST-3PL whoever carry.PST.3SG=3PL in that_same
madrese.kār.dāneš bači-yal bord=ešun.
technical.school child-PL carry.PST.3SG=3PL
[they went,] someone (lit. whoever) took them into that same technical school, the young people took them.
- (89) *hālā harki bord=ešun bord=ešun učo ...*
now whoever carry.PST.3SG=3PL carry.PST.3SG=3PL there
Well, whoever took them, took them there ...
- (90) *tā haf ruz=am xevar ne-dāšt-im ez=ešun.*
until seven day=ADD news NEG-have.PST-1PL from=3PL
For seven days we still had no news of them.
- (91) *hič na dunest-im gā koja-n=e na či ...*
any not know.PST-1PL cow where-n=3SG not what
We didn't know where the cows were, or anything (lit. any of us neither knew where the cows are, nor anything) ...
- (92) *hālā ke bāvā hamzi ovay*
now COMPL grandfather come.PST.3SG come.PST.3SG
māmā kemutar rah jelov=š ogo
grandmother Kemutar go.PST.3SG in_front=3SG say.PST.3SG
pa hamzi mand-i hamičo, bači-yal=om xu bi
so Hamzi stay.PST-2SG right_here child-PL=1SG good be.PST.3SG
na-mord-en!
NEG-die.PST-3PL
As Grandpa Hamzi arrived Grandma Kemutar went up to him and said to him, “So Hamzi, you stayed right here?! Our (lit. my) children were fine – they didn’t die!”
- (93) *rahd-en zir ogo na, hiči=šun*
go.PST-3PL down say.PST.3SG no something=3PL
na-vavi!
NEG-become.PST.3SG
They went down, he said, “No! Nothing happened!”

- (94) *bāvā hamzi ez pele-ā er-e*
 grandfather Hamzi from stair-PL go.NPST-3SG
er-e hamučo
 go.NPST-3SG right_there
ogo na, b-iyo hamičo be-nešin-im.
 say.PST.3SG no IMPV-come.IMPV.2SG right_here IMPV-sit.NPST-1PL
 Grandfather Hamzi goes to the steps, he goes right over here, he said, “No!
 Come and sit right here”.
- (95) *nešast ba=am hamučo xavsi mi tuv-ā*
 sit.PST.3SG father=1SG right_there sleep.PST.3SG in room-PL
ogo hiči=šun ni.
 say.PST.3SG anything=3PL be.PST.NEG.3SG
 My father sat right there, he lay down in the house, he said, “Everything is
 okay (lit. their anything is not)”.

Abbreviations

1	first person	NEG	negative
2	second person	NPST	non-past
3	third person	NUM	numeral classifier
AA	Anonby & Asadi (see References)	O/OBJ	object
ADD	additive marker	PART	participle
adj.	adjective	PF	perfect
ADJZR	adjectivizer	PL	plural
AT	Anonby & Taheri-Ardali (see References)	PROX	proximal
AUX	auxiliary	PST	past
B.	Bakhtiari	SBJNC	subjunctive
C	consonant	SG	singular
C&B	Chahar Mahal va Bakhtiari	SPEC	specifier
CMPV	comparative	tr.	transitive
COMPL	complementizer	V	vowel; verb
cont.	continous	Ā	long vowel
DEF	definite	VOC	vocative
DIST	distal	~	varies between
EZ	<i>ezāfe</i> (associative morpheme)	-	morpheme boundary
IMPFV	imperfective	=	clitic attachment
IMPV	imperative	'	stress/accent
INDEF	indefinite	.	syllable boundary
n.	noun		

References

- Abbott, Barbara. 2004. Definiteness and indefiniteness. In Laurence R. Horn & Gregory Ward (eds.), *Handbook of pragmatics*, 122–149. Oxford: Blackwell.
- Abdullaeva, Firuza. 2009. Zhukovskii, Valentin Alekseevich. *Encyclopædia Iranica*. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/zhukovskii-valentin-alekseevich> (accessed 25 September 2017).
- Afsar Bakhtiyāri, Dārāb. 1965/1344. *Montakhabāti az aš'ār-e šā'er-e farzāneh Dārāb Afsar Bakhtiāri* [A selection from the poems of the eminent poet Dārāb Afsar Bakhtiyāri], ed. by Abd-ol-Hossein Sepantā. Tehran: Mash'al.
- Amanollahi, Sekandar. 1991/1370. *Qowm-e Lor* [The Luri people]. Tehran: Āgāh.
- Anonby, Erik. 2003. Update on Luri: How many languages? *Journal of the Royal Asiatic Society* 13(2). 171–97.
- Anonby, Erik. 2012. Sociolinguistic status of Lori. *Encyclopædia Iranica*. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/lori-language-ii> (accessed 25 September 2017).
- Anonby, Erik. 2016. Mapping the languages of western Iran's Zagros Mountains. Network Meeting of the Alexander von Humboldt Foundation, University of Freiburg, 5–7 October, 2016.
- Anonby, Erik. 2017. The languages of Iran: A working classification. In AT 2015–18.
- (AA 2014=) Anonby, Erik & Ashraf Asadi. 2014. *Bakhtiari studies: Phonology, text, lexicon*. Studia Iranica Upsaliensia 24. Uppsala: Acta Universitatis Upsaliensis.
- (AA forthcoming=) Anonby, Erik & Ashraf Asadi. (forthcoming). *Bakhtiari studies: Bakhtiari orthography*. Studia Iranica Upsaliensia. Uppsala: Acta Universitatis Upsaliensis.
- (AT 2015–18=) Anonby, Erik & Mortaza Taheri-Ardali (eds.). 2015–18. *Atlas of the languages of Iran*. Ottawa: GCRC (Geomatics and Cartographic Research Centre), Carleton University. <http://iranatlas.net> (accessed 25 September 2017).
- Anonby, Erik & Mortaza Taheri-Ardali. 2017. Chahar Mahal va Bakhtiari Province in the *Atlas of the Languages of Iran*. Paper presented at the 1st North American Conference in Iranian Linguistics (NACIL1), Stony Brook University, New York, 28–30 April, 2017.
- Dāvari, Bahram. 1964/1343. *Zarb-ol-masalhā-ye Bakhtiāri* [A collection of Bakhtiari proverbs]. Tehran: Tahuri.
- Digard, Jean-Pierre. 1988. Bakhtiāri tribe: Ethnography. *Encyclopædia Iranica* 3(5). 553–560. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/bakhtiari-tribe#pt1> (accessed 25 September 2017).
- Digard, Jean-Pierre. 1992. Clothing of the Bakhtiāris and other Lori speaking tribes. *Encyclopædia Iranica* 5.8, 852–854. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/clothing-xxv> (accessed 25 September 2017).
- Forutan, Fariborz. 1996/1375. *Ganjine-ye masalhā-ye Bakhtiyāri* [A treasury of Bakhtiari proverbs]. Tehran: Enteshārāt-e Khuzestān.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A Construction Grammar approach*. Berlin & Boston: De Gruyter Mouton.
- Haig, Geoffrey. 2017. Post-predicate arguments in Iranian languages. Paper presented at the 7th International Conference on Iranian Linguistics (ICIL7), Moscow State University, 28–30 August, 2017.

- Huddleston, Rodney & Geoffrey K. Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Ittig, Annette. 1988. Bakhtiari tribe: Bakhtiari carpets. *Encyclopædia Iranica* 3(5). 553–560. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/baktiari-tribe#pt3> (accessed 25 September 2017).
- Kerimova, Aza A. 1982. Lurskie i Bakhtiyarskie dialekty [Luri and Bakhtiari dialects]. In Vera S. Rastorgueva (ed.), *Osnovy iranskogo iazykoznaniiia iazyki III. Novoiranskie iazyki: zapadnaia gruppa, prikasiiskie iazyki* [Fundamentals of Iranian languages III: New Iranian languages – Western group, Caspian languages], 287–315. Moscow: Nauka.
- Khosravi, Abdol’ali. 1996/1375. *Guyesh-e Bakhtiari: Ketāb-e chahārom* [Bakhtiari dialect: Fourth book]. Esfahan: Nashr-e Ghazal.
- Khosraviniyā, Javād. 2010/1389. *Shāhnāme Bakhtiyāri* [Bakhtiari Shahnameh]. Esfahan: Shāhsavāri.
- Khosraviniyā, Javād. 2011/1390. *Gowgari* [Brotherhood]. Shahr-e Kord: Niyosheh.
- Kiyāni Haft Lang, Kiyānush. 2000/1378. *Zarb-ol-masalhā-ye Bakhtiāri* [Bakhtiari proverb], 2nd edn. Ahvaz: Enteshārat-e Khuzestān.
- Lazard, Gilbert. 1985. L’inférentiel ou passé distancié en persan. *Studia Iranica* 14(1). 27–42.
- Lazard, Gilbert. 2001. On the grammaticalization of evidentiality. *Journal of Pragmatics* 33(3), 359–367.
- Lorimer, David L. R. 1922. *The phonology of the Bakhtiari, Badakhshani, and Madaglashti dialects of modern Persian, with vocabularies*. London: Royal Asiatic Society.
- Lorimer, David L. R. 1930. A Bakhtiari prose text. *Journal of the Royal Asiatic Society*, April 1930, 347–364.
- Lorimer, David L. R. 1954/1955/1963. The popular verse of the Bakhtiari of southwestern Persia I/II/III. *Bulletin of the School of Oriental and African Studies* 16/17/26. 542–555/92–110/55–68.
- Lorimer, David L. R. 1964. A Bakhtiari Persian text. In Georges Redard (ed.), *Indo-Iranica: Mélanges présentés à Georg Morgenstierne*, 129–133. Wiesbaden: Harrassowitz.
- Lorimer, David L. R. & E. O. Lorimer. 1919. *Persian tales, written down for the first time in the original Kermani and Bakhtiari*. London: Macmillan.
- MacKinnon, Colin. 2011. Lori dialects. *Encyclopædia Iranica* 18. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/lori-dialects> (accessed 25 September 2017).
- Madadi, Zohrāb. 1996/1375. *Vāzhehnāmeh-ye zabān-e Bakhtiāri* [Lexicon of the Bakhtiari language]. Tehran & Esfahan: Enteshārat-e Āyat.
- Mann, Oskar. 1910. *Die Mundarten der Lur Stämme im südwestlichen Persien*, Berlin: Georg Reimer.
- Perry, John R. 1996. Persian in the Safavid period: Sketch for an Etat de langue. In Charles Melville (ed.), *Safavid Persia: The history and politics of an Islamic society*, 269–283. London: I.B. Tauris.
- Ra’isi, Dārāb. 1996/1375. *Golbāng-e Bakhtiyāri* [Cry of the Bakhtiari], 3rd edn. Shahr-e Kord: Enteshārat-e Il.
- Schreiber, Laurentia, Mortaza Taheri-Ardali, Geoffrey Haig & Erik Anonby. 2017. Contact-induced change in Irano-Turkic morphosyntax. Paper presented at the Conference

- on Language Contact and Language Change in Western Asia, Goethe University Frankfurt, 10–12 March, 2017.
- Shirin Bayān, Sa' id. 2007/1386. *Yek eyl, yek sedā* [One tribe, one voice]. Tehran: Serāj-e Nur.
- Stilo, Donald. 2010. Ditransitive constructions in Vafsi: A corpus-based study. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook*, 243–276. Berlin & Boston: De Gruyter Mouton.
- Tāheri, Esfandyār. 2010/1389. *Guyesh-e Bakhtiāri-ye Kuhrang* [The Bakhtiari dialect of Kuhrang]. Tehran: IHCS Press.
- Taheri-Ardali, Mortaza. 2017. Salient linguistic features and isoglosses in Chahar Mahal va Bakhtiari Province, Iran. Paper presented at the 1st North American Conference in Iranian Linguistics (NACIL1), Stony Brook University, New York, 28–30 April, 2017.
- Taheri-Ardali, Mortaza. 2015. Guyesh-e Bakhtiāri: zavāl yā baghā? Negāhi digar be kārbord-e guyesh-e Bakhtiāri, guneh-ye Ardal [Bakhtiari dialect: decline or survival? A re-examination of the usage of the Bakhtiari dialect, Ardal variety]. Paper presented at the International Conference on Linguistic Heritage, Tehran, 20–21 October, 2015.
- Taheri-Ardali, Mortaza & Erik Anonby. 2017. Language distribution in Chahar Mahal va Bakhtiari Province of Iran. Paper presented at the 1st International Conference on Languages, Dialects and Linguistics (ICCILDL1), Pazhoheshgaran Andishmand Institute, Ahvaz, Iran, 2–3 February, 2017.
- Wal Anonby, Christina van der. 2011. Remarks on the remarkable: Mirativity in Kumzari. Paper presented at the 4th International Conference on Iranian Linguistics (ICIL4), Uppsala University, Uppsala, 17–19 June, 2011.
- Windfuhr, Gernot. 1988. Bakhtiāri tribe: The Bakhtiāri dialect. *Encyclopædia Iranica* 3(5). 553–560. New York: Center for Iranian Studies, Columbia University. <http://wwwiranicaonline.org/articles/baktiari-tribe#pt2> (accessed 25 September 2017).
- Windfuhr, Gernot. 2009. Dialectology and topics. In Gernot Windfuhr (ed.), *The Iranian languages*, 5–42. London & New York: Routledge.
- Zhukovskij, Valentin A. 1923. *Materialy dlya izucheniya persidskikh narechij, pt. 3: Narechie bakhtiyarov chekharleng i kheftleng* [Materials for the study of Persian dialects, part 3: The dialect of the Chahārlang and Haftlang Bakhtiari], ed. by S. F. Ol'denburg. Petrograd.
- Zhukovskij, Valentin A. 1923/2017. *Mavāddi barā-ye motāle'eh-ye guyesh-e Bakhtiāri* [Materials for the study of the Bakhtiari dialect], transl. and ed. by Maryam Shafaghi and Seyyed Mehdi Dadras. Tehran: Allameh Tabataba'i University Press.
- Zolfaghari, Sima. 2017. *The Bakhtiaris: An anthropological-linguistic lexical study of Haft Lang nomads of southwest Iran*. Leiden: Leiden University dissertation.

4.4. The Neo-Aramaic dialects of western Iran

Geoffrey Khan

1. Genetic affiliation

The Neo-Aramaic dialects that are the subject of this chapter belong to two sub-groups. They include:

- (i) dialects belonging to the North-Eastern Neo-Aramaic (NENA) subgroup, which were spoken in the Kordestan and Kermanshah provinces of western Iran until the middle of the twentieth century.¹
- (ii) The neo-Mandaic subgroup of dialects, which were spoken further south in the towns of Ahvāz and Khorramshahr in Khuzestan province of western Iran.

The NENA dialect area includes also the region of northern Iraq (chapter 3.4) and the region of southeastern Turkey and northwestern Iran (chapter 2.5). The NENA dialects of these latter two regions exhibit considerable diversity, as has been shown in the chapters devoted to them. The NENA dialects of the region described in this chapter, by contrast, exhibit very little diversity. The main split is between Jewish and Christian varieties of NENA. The Jewish variety consists of a cluster of dialects spoken by Jewish communities in various localities in an area that includes Sainqala, Bokan, Sāqəz on its northern border, Sanandaj in the centre, Bijar on the eastern border, and in the south Kerend and Qasr-e Širin. The Christians variety consists of a single dialect spoken by Christians in the town of Sanandaj.

The Jewish cluster of dialects is remarkably uniform and only minor differences are found among the dialects of the aforementioned places where the dialects were spoken. Most of the data on this variety of NENA for this chapter are taken from the Jewish dialect of Sanandaj (henceforth referred to as J. Sanandaj),² which has been described in detail in Khan (2009). Studies on other Jewish dialects of this cluster include those by Israeli (1998) on the dialect of Sāqəz and by Hopkins (2002) on Kerend. The Jewish dialects of the region belong to the so-called trans-Zab subgroup of Jewish NENA. Within trans-Zab they are most closely related to the Jewish dialect of Sulemaniyya in northeastern Iraq (chapter 3.4 and Khan 2004). Their relationship to the Jewish trans-Zab cluster of dialects of the West Azerbaijan province of Iran in the Urmi region is more distant to the extent that

¹ The term North-Eastern Neo-Aramaic was coined by Hoberman (1988: 557).

² As in chapters 2.5 and 3.4, Jewish dialects are distinguished from Christian dialects in this chapter by the prefixing J. and C. respectively to the names of the locations where they were spoken.

speakers of dialects from the western Iran cluster had difficulties communicating with Jews from Urmi.

The Christian NENA dialect of Sanandaj (henceforth referred to C. Sanandaj) is, likewise, very similar to the Christian dialect of Sulemaniyya (chapter 3.4) but substantially different from the Christian dialects spoken in the Urmi region of Iran. Grammatical and lexical studies on the C. Sanandaj include Panoussi (1990, 1991), Heinrichs (2002) and Kalin (2014). Short extracts of texts in the dialect can be found in Panoussi (1990: 120–128) and Macuch and Panoussi (1974: 39).³ These authors refer to the dialect as the Senaya dialect, from the Kurdish name of the town ‘Sena’. A brief overview of the NENA dialect situation in Iran in general can be found in Hopkins (1999).

Neo-Mandaic is spoken by Mandaean in southwestern Iran. This exists in two known varieties originally spoken in the towns of Ahvāz and Khorramshahr respectively (Macuch 1965, 1989, 1993; Häberl 2009, 2011; Mutzaifi 2014). The native name of this spoken language is *ratnā*, derived from a verbal root meaning ‘to whisper or mutter’.

2. History of the speech communities

Very little is documented concerning the history of the NENA-speaking communities before the twentieth century. The lack of diversity in the Jewish cluster of dialects suggests that the communities who spoke them migrated in a single wave into the region in relatively recent centuries. The ancient heartland of NENA must have been in what is now northern Iraq and southeastern Turkey, where there is great dialectal diversity. The close relationship of the Jewish dialects of western Iran with the J. Sulemaniyya dialect points to northeastern Iraq as the origin of the migration. The isolated Christian dialect spoken in Sanandaj must, likewise, have been the result of migration. It is closely related to the Christian dialect of Sulemaniyya and there were family relationships between these two communities in living memory of speakers. It is likely, therefore, that the Christians of Sanandaj were originally migrants from the region of Sulemaniyya in Iraq.

We know that some of the Jewish communities who settled in the towns of the region originally lived in surrounding villages. The Jews of Sanandaj, for example, moved into the town after its foundation in the 17th century from a village known as Qal‘at Ḥasan’ābād (Khan 2009: 1).

The Christians of Sanandaj belong to the Chaldean Church. In the 19th century several Christian families moved to Qazvīn, where their speech developed the

³ Data on C. Sanandaj cited in this chapter are mainly taken from the publications of Panoussi or from personal communications from him, which I acknowledge with gratitude.

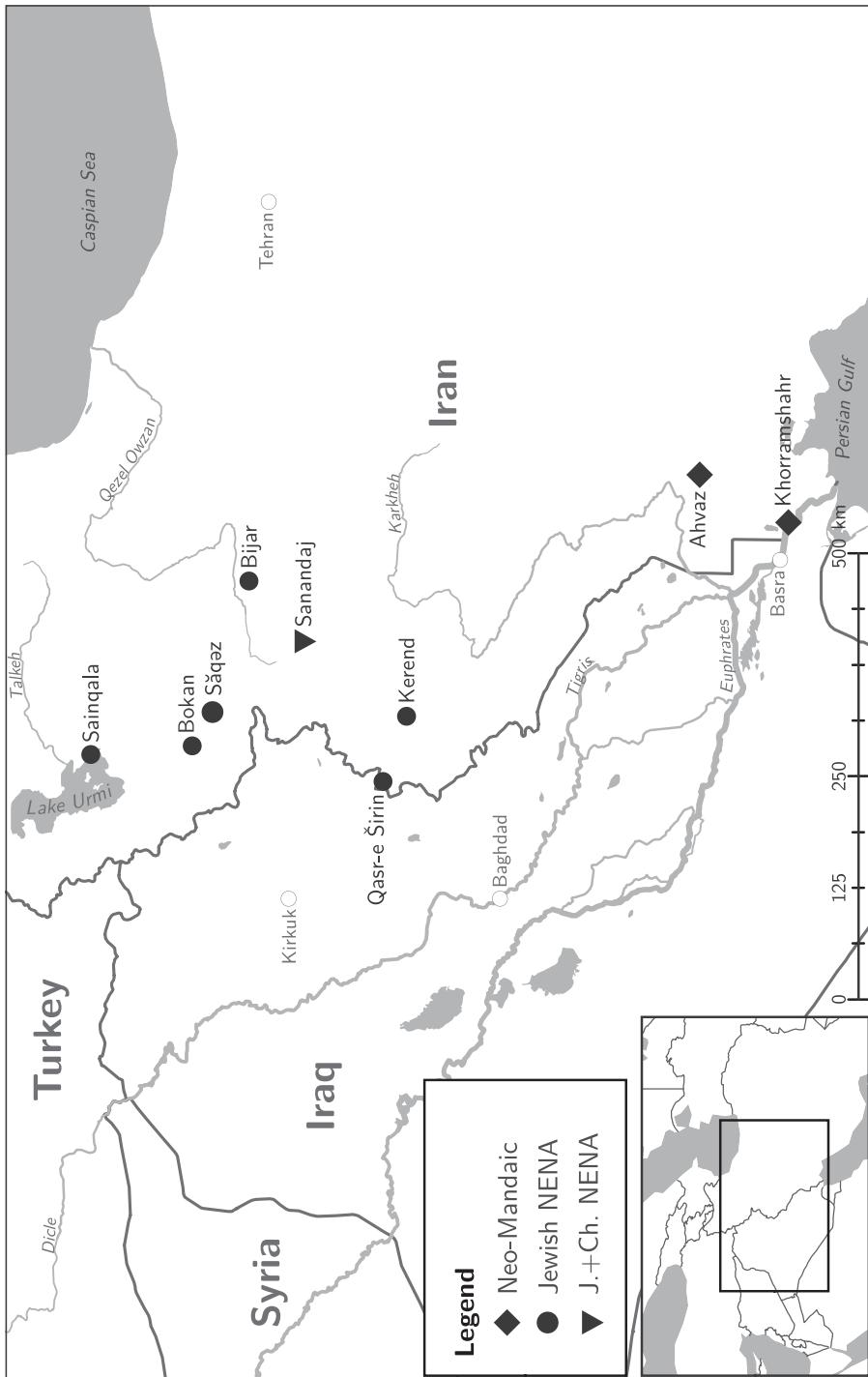


Figure 1: Locations of NENA and Neo-Mandaic varieties covered in this chapter

distinctive trait of the realization of /w/ as /v/ under the influence of Persian (Heinrichs 2002: 238). In the middle of the twentieth century the Chaldean diocese of Sanandaj moved to Tehran and the Christian Neo-Aramaic speakers moved with it.

Neo-Mandaic is spoken by the Mandaeans, who follow a religion that is a descendant of a pre-Islamic Gnostic sect. The traditional homeland of the Mandaean community is the south of Iraq and the adjacent Khuzestān province of southwest Iran. They are known in Iraq and Iran as ‘Sabians’ (Arabic *ṣābi’ūn*, colloquial *ṣubba*), who are one of the ‘peoples of the book’ (*’ahl al-kitāb*) recognized in Islam.

Neo-Mandaic appears to have ceased to be the spoken language of the Mandaeans of Iraq by the beginning of the 19th century. There are references to a few speakers in Iraq in the 20th century, but these seem to be of Iranian origin (Häberl 2009: 36–37). After the first Gulf War in 1991 the Iraqi Mandaean community was displaced from their homes in southern Iraq by the army of Saddam Hussein.

Up until the 19th century Neo-Mandaic was spoken in a variety of localities in the Khuzestān region. The Mandaeans subsequently came to be concentrated in Khorramshahr and Ahvāz, where two distinct varieties of the language survived until modern times. During the Iranian revolution in 1979 and in subsequent conflicts with Iraq, Khorramshahr was largely destroyed and abandoned by its inhabitants, including the Mandaean community. Within Iran the language seems now to be spoken only in Ahvāz (Mutzafi 2014: 1–5).

3. Current status of the speech communities

In 1952 many NENA-speaking Jews from the region emigrated to the newly founded State of Israel. Over the subsequent two decades there was a gradual emigration of the Jews either to Tehran or abroad, mostly to Israel. After the Iranian Revolution in 1979 most of the remaining Jews left the region, the majority settling in Los Angeles in the USA and the remainder in Israel or Europe. Today only about half a dozen elderly Jews are reported to be still living in the town.

After the Christian NENA-speaking community moved *en masse* from Sanandaj to Tehran, they gradually left Iran and settled abroad. The majority of the migrants have settled in the Los Angeles area of the USA.

As a result of these migrations and the disintegration of the NENA speech communities, the NENA dialects that were spoken in western Iran are now highly endangered.

There are numerous Mandaeans living in the urban centres of Iraq and in communities that have settled outside of the Middle East, especially in Sweden, Australia and the USA. The vast majority, however, do not speak neo-Mandaic. The number of competent speakers of the language is rapidly dwindling. Häberl (2009: 8) estimated there to be around 100–200 elderly speakers, most of whom are living in Iran. Neo-Mandaic, therefore, will inevitably become extinct within the next few years.

4. Sociolinguistic situation

The Muslims in the region speak varieties of Sorani Kurdish. Persian is the official language in schools and government administration. Most Jews and Christians of the region were able to communicate fluently in Kurdish and Persian (Farsi) as well as Neo-Aramaic. This resulted in extensive influence of these contact languages on the Neo-Aramaic dialects in all levels of the grammar and in the lexicon.

The Mandaean in the Iranian province of Khuzestān are trilingual. In addition to neo-Mandaic they speak the local dialects of Arabic, which constitute the vernacular of much of the Muslim population of the area, and also Persian, which is the official language and language of education. Some of the speakers use an adapted form of the Classical Mandaic script to write down the vernacular language.

5. Phonology

5.1. *bgdkpt* consonants and *q*

In earlier Aramaic the stop consonants *bgdkpt* developed fricative allophones after vowels, which can be represented *[b], *[g], *[d], *[k], *[p̪], *[t̪].

5.1.1. NENA dialects

In the NENA dialects of the region the fricatives *[b], *[g], *[k], and *[p̪] have reflexes that are common to the NENA subgroup as a whole (see chapter 2.5, §5.1 and chapter 3.4, §5.1):

*[b] > /w/	J. Sanandaj: <i>dənwe</i> (< * <i>dənbē</i>) ‘flies’	C. Sanandaj: <i>gānawa</i> (< * <i>gannābā</i>) ‘thief’
------------	--	---

In the variety of C. Sanandaj spoken in Qazvīn this reflex is pronounced /v/.

*[p̪] > /p/	J. Sanandaj: <i>kepa</i> (< * <i>kēpā</i>) ‘stone’	C. Sanandaj: <i>napel</i> (< * <i>nāpēl</i>) ‘he falls’
*[k] > /x/	J. Sanandaj <i>băxe</i> (< * <i>băkē</i>) ‘he weeps’	C. Sanandaj: <i>baxe</i>

The velar fricative *[g] has in most cases been weakened to zero, e. g.

*[g]	J. Sanandaj: <i>śrata</i> (< * <i>śrāḡtā</i>) ‘lamp’	C. Sanandaj: <i>čraya</i> (< * <i>śrāḡā</i>) ‘lamp’
------	--	---

This process of weakening of the velar fricative *g involved an intermediate stage in which the velar shifted to the pharyngeal * (Tsereteli 1990). In J. Sanandaj the

pharyngeal is preserved in the word *lo ‘á* ‘inside’, the development of which can be reconstructed as *lo ‘á* < **l-*‘oya (by metathesis, this being the form of the word in J. Amedia) < **l-ğoya*. The preservation of the pharyngeal would have been conditioned by a pharyngealized pronunciation of the word, no doubt facilitated by the /*l*/*/, though this has now been lost. The word is still pronounced with suprasegmental pharyngealization in J. Urmī (+*lwa* [l^fw^fq^f]).*

Jewish and Christian NENA of the region differ with regard to the reflexes of the interdental fricatives *[d] and *[t]. In the Jewish cluster of dialects the reflex of both is generally the lateral /*l*/*/, as in other trans-Zab Jewish dialects:*

- | | |
|-------------|---|
| J. Sanandaj | <i>bela</i> (< * <i>baytā</i>) ‘house’ |
| | <i>qlila</i> (< * <i>qdila</i>) ‘key’ |

The lateral /*l*/*/ occurs in a number of words where the cluster of Jewish trans-Zab dialects in northwestern Iran have /*d* rather than the normal reflex /*l*/*/:**

J. Urmī	J. Sanandaj	
<i>ida</i>	<i>’ila</i>	< * <i>’idā</i> ‘hand’
<i>od</i>	<i>’ol</i>	< * <i>’ābəd</i> ‘he does’

In a few words in the Jewish dialects the reflex of the unvoiced interdental **t* is the unvoiced pharyngeal fricative /*h*/*/, e. g.*

- | | |
|-------------|--|
| J. Sanandaj | <i>’ahra</i> (< * <i>’atrā</i>) ‘town’ |
| | <i>təlħa</i> (< * <i>tlātā</i>) ‘three’ |
| | <i>lāħmal</i> (< * <i>lā-timmal</i>) ‘the day before yesterday’ |
| | <i>nahale</i> (< * <i>nāħatā</i>) ‘ears’ |

In such words the **t* was originally weakened to **h* and this subsequently shifted to a pharyngeal /*h*/*/ . Such words must have been originally pronounced with suprasegmental pharyngealization, originating no doubt from the consonants *r*, *l* or *m*. The pharyngealization was subsequently lost as a suprasegmental feature but left a vestige in the pharyngeal segment /*h*/*/: *’ahra* < **’aħrā* < **’atrā*. The intermediate stage of development /*h*/*/ and pharyngealization is found in J. Urmī: ⁺*ahra*, ⁺*nahale*.***

In some plural forms of nouns in the Jewish dialects the reflex of **t* is zero, e. g.

- | | |
|-------------|--|
| J. Sanandaj | <i>malāwāe</i> (< * <i>mātawātā</i>) ‘villages’ |
|-------------|--|

An original /*d*/*/ in post-vocalic position in the Jewish dialects sometimes shifts to the fricative /*z*/*/:**

J. Sanandaj	<i>koza</i> ‘liver’	cf. J. Urmī <i>koda</i>
	<i>guza</i> ‘wall’	cf. J. Urmī <i>guda</i>

The articulation of the consonant has been further weakened in a few cases to zero, e. g.

J. Sanandaj	<i>xar</i> ‘He becomes’	cf. J. Urmi <i>xadər</i>
	<i>šar</i> ‘He sends’	cf. J. Urmi <i>šadər</i>

In C. Sanandaj the reflexes of *[d] and *[t] are /d/ and /s/ respectively, as in C. Sulemaniyya in Iraq, e. g.

C. Sanandaj	<i>ida</i> (< *'idā) ‘hand’
	<i>besa</i> (< *baytā) ‘house’

The /s/ may become pharyngealized in environments that are liable to pharyngealization, such as words containing the sonorant /r/:

C. Sanandaj	<i>baṣra</i> (< *batrā) ‘behind’
-------------	----------------------------------

In the Jewish dialects /q/ is normally realized as an unvoiced uvular stop, e. g. *baqá* [ba:’qa] ‘to’. After a vowel or /w/, it is occasionally realized as an unvoiced uvular fricative (Khan 2009: 21), e. g.

J. Sanandaj		
<i>qoqé</i>	[qo:’χe:]	‘pots’
šəwqá-y	[ʃɪf’χaj]	‘He has left’.

5.1.2. Neo-Mandaic

Neo-Mandaic has preserved all of the fricative *bgdkpt* consonants except fricative *d, which has shifted to a stop /d/ (Macuch 1965: 32–40; Häberl 2009: 48–65). The examples below are taken from Häberl’s description of the Khorramshahr dialect:

* <i>b</i>	[v]	<i>gabrā</i> ['gævrə]⁴ ‘man’, [w] in the environment of back rounded vowels: <i>əbod</i> [ə'wod] ‘do! (ms.)’
* <i>g</i>	[ɣ]	<i>palḡā</i> ['palɣə] ‘split’, <i>logrā</i> ['loyərə] ‘leg’
* <i>d</i>	[d]	<i>idā</i> ['i:də] ‘hand’
* <i>k</i>	[χ]	<i>əkal</i> [a'χal] ‘he ate’
* <i>p</i>	[f]	<i>nəpāq</i> [nə'faq] ‘he went out’
* <i>t</i>	[θ]	<i>bietā</i> ['biɛθə] ‘house’, <i>hātā</i> [hɔ:θə] ‘sister’

⁴ Following the practice of Häberl in his publications, in the section on phonology examples from Neo-Mandaic are given in a phonemic transcription in italics and in a phonetic transcription in IPA characters in roman enclosed in square brackets.

The shift of the original **d* to the stop *d* in Neo-Mandaic seems to be a relatively recent development. A glossary of the spoken language written in Mandaic script in the 17th century indicates the existence of the fricative pronunciation of the letter in many native Aramaic words by a diacritical dot (Borghero 2004: 70–71).

Before a syllable beginning with the sonorant [l], the unvoiced interdental [θ] loses its oral articulation and shifts to [h], e. g.

<i>ehle</i>	[‘eħle]	‘I have’	< *eθ-le
-------------	---------	----------	----------

In addition to the inherited fricative *bgdkpt* consonants, Neo-Mandaic also exhibits the fricativization of the inherited voiced stops /b/, /g/, /d/ to the voiceless fricatives /f/, /χ/ and /θ/ respectively before voiceless stops (Häberl 2009: 89), e. g.

<i>rabbā</i> ['ræbbə]	‘great’	<i>rap̄tar</i> ['ræftær]	‘greater’
-----------------------	---------	--------------------------	-----------

In Neo-Mandaic the voiceless uvular stop /q/ is often realized as voiced uvular fricative [χ] after a vowel in a stressed syllable and in word-initial position it is occasionally realized as a voiced uvular stop (Häberl 2009: 57), e. g.

<i>láqben</i>	[‘laχ.ben]	‘we don’t want’
<i>qēhazelli</i>	[ga.‘zel.li]	‘they see him’

5.2. Laryngeals and pharyngeals

5.2.1. NENA dialects

In the Jewish NENA cluster of dialects of the region, an initial laryngeal stop *’ in some words shifts to a laryngeal fricative /h/, e. g.

J. Sanandaj	<i>hāmər</i>	‘(that) he says’	< ‘āmər
	<i>hezəl</i>	‘(that) he goes’	< ‘ezəl

As generally in NENA, the unvoiced pharyngeal fricative **h* has shifted to the velar fricative /χ/ in both Jewish and Christian NENA of the region, e. g.

J. Sanandaj	<i>xmara</i> (< *hmārā)	‘ass’	<i>qamxa</i> (< *qamhā)	‘flour’
-------------	-------------------------	-------	-------------------------	---------

C. Sanandaj	<i>xmara</i>	<i>qamxa</i>
-------------	--------------	--------------

The original pharyngeal has been preserved in some words and verbal roots of Aramaic stock that contain /q/ or a historical pharyngealized consonant, e. g.

J. Sanandaj	<i>ħānəq</i>	‘he drowns’
C. Sanandaj	<i>rahúqa</i>	‘distant’

In C. Sanandaj **h* is preserved in some words relating to religion which are taken from Classical Syriac, e. g. *mšiha* ‘Messiah’.

In word initial position the reflex of an historical voiced pharyngeal fricative *‘ is normally the laryngeal stop /χ/, e. g.

- | | |
|-------------|--|
| J. Sanandaj | <i>'ela</i> (< * <i>ēdā</i>) ‘festival’ |
| C. Sanandaj | <i>'eda</i> |

In word-internal or word-final position the voiced pharyngeal *‘ has been weakened to zero in most cases in both Jewish and Christian NENA, e. g.

- | | |
|-------------|---|
| J. Sanandaj | <i>beta</i> (< * <i>bētā</i>) ‘egg’, <i>šame</i> (< * <i>šāma</i>) ‘he hears’ |
| C. Sanandaj | <i>tēna</i> (< * <i>ṭa'nā</i>) ‘load’; <i>šame</i> (< * <i>šāma</i>) |

The pharyngeal *‘ has been retained in some words of Aramaic stock that contain /q/ or a consonant that was historically pharyngealized, e. g.

- | | |
|-------------|---|
| J. Sanandaj | <i>'aqəwra</i> (< * <i>aqəbrā</i>) ‘scorpion’; <i>ṭa'na</i> (< * <i>ṭa'nā</i>) ‘load’ |
| C. Sanandaj | <i>'arəq</i> (< * <i>ārəq</i>) ‘he flees’; <i>'arəṭ</i> (< * <i>ārəṭ</i>) ‘he farts’ |

In some cases its reflex is a laryngeal /χ/ in such contexts:

- | | |
|-------------|--|
| C. Sanandaj | <i>ṣa'or</i> (< * <i>ṣā'or</i>) ‘he upbraids’ |
|-------------|--|

Pharyngeals are also preserved in the Jewish dialects in environments where a consonant adjacent to the pharyngeal became pharyngealized at some stage of development. This is found especially in the environment of the sonorant consonants /m/, /b/, /l/ and /r/, e. g.

- | | |
|-------------|---|
| J. Sanandaj | <i>dəm'e</i> (< * <i>dəm'e</i> , < * <i>dəm'e</i>) ‘tears’ |
| | <i>gul'a</i> (< * <i>gul'a</i> < * <i>gul'ā</i>) ‘kernel of fruit’ |
| | <i>za'ra</i> (< * <i>za'ra</i> , < * <i>s'ārā</i>) ‘barley’ |

5.2.2. Neo-Mandaic

The original Aramaic pharyngeals have been lost in Neo-Mandaic. The normal reflex of **h* is /h/ and that of *‘ is zero.

- | | | |
|-----------------------|---------|--------------------|
| <i>hamšā</i> [ˈhæmʃā] | ‘five’ | < * <i>hamšā</i> |
| <i>ālmā</i> [ˈvlmā] | ‘world’ | < * <i>ālmā</i> |
| <i>ārbīn</i> [ɔɪ'bɪn] | ‘forty’ | < * <i>'arb īn</i> |

When /h/ closes a syllable in word-medial position, it is often strengthened to [x] in order to prevent its elision, e. g.

- | | |
|-----------------------------------|-------------|
| Khorramshahr (Häberl 2009: 59) | |
| <i>qəhazenkən</i> [qax. 'zin:χon] | ‘I see you’ |

The speakers of Neo-Mandaic from Khuzestān are today all bilingual in Arabic and their Neo-Mandaic contains loanwords from Arabic containing the Arabic pharyngeal consonants. In the pronunciation of these loanwords the pharyngeals are either retained as they are articulated in Arabic or they are weakened to laryngeals, most likely under the influence of Persian (Häberl 2009: 58–59), e. g.

Khorramshahr

<i>ḥākem</i>	[ˈhɒ:kəm]	‘governor’	< Arabic <i>ḥākim</i> ‘judge’
<i>hašiš</i>	[hæ:ˈʃiʃ]	‘hashish’	< Arabic <i>hašiš</i> ‘grass’
<i>sun ‘a</i>	[‘sʌn̩fɑː]	‘good’	< Arabic <i>sun</i> ‘benefit’
<i>jami ‘a</i>	[dʒe'mi:ʔa]	‘all’	< Arabic <i>jamī</i> ‘all’

5.3. Pharyngealized consonants

5.3.1. NENA dialects

C. Sanandaj preserves the pharyngealized consonants /ṣ/ and /ṭ/, e. g. *ṣale* ‘he prays’, *ṭaše* ‘he hides’.

In the Jewish dialects the original pharyngealization of these consonants is generally weakened. There is, therefore, minimal spreading of pharyngealization in the word (Khan 2009: 17–18, 24), e. g.

J. Sanandaj *tura* [tʰu:’ra] ‘mountain’, *xasa* [xa:’sa] ‘back’.

If, however, there was an original pharyngeal fricative *‘ in the word, this is retained in words containing the originally pharyngealized consonants /ṣ/ and /ṭ/, e. g.

J. Sanandaj *ṭa ‘na* [tʰaʃ‘na] ‘load’ (< *ṭa ‘nā)

This process of development seems to be due to convergence with the behaviour of pharyngealization in the Kurdish dialects described by Margaret Kahn (1976: 49–52). In Kahn’s analysis a pharyngealized consonant in Kurdish is the result of the presence of an underlying pharyngeal segment in the word. Such words can also be realized with the pharyngeal on the phonetic surface in place of the pharyngealized consonant.

The Jewish dialects of the region do, however, have a pharyngealized sonorant /ɻ/. It occurs in loanwords from Kurdish, e. g.

J. Sanandaj *pule* ‘money’, *qali* ‘carpet’, *zuxāl* ‘coals’

It is found also in a number of words of NENA origin. In such cases it corresponds to original /r/, e. g.

J. Sanandaj *pšila* (< **pšira*) ‘melted’; *jołe* (< **jore*) ‘urine’

5.3.2. Neo-Mandaic

As in the Jewish NENA dialects of the region, in Neo-Mandaic the pharyngealization of historical **t* and **s* is often weakened. This weakening is particularly common when these consonants are in contact with a following non-pharyngealized consonant (Häberl 2009: 53–54), e. g.

Khorramshahr

<i>tāb</i>	[t̪ɔv]	'good'
<i>batluktā</i>	[bat'lʊχtɔ]	'misfortune'
<i>səbyi</i>	[s̪evjɪ]	'he baptized him'
<i>mīṣrā</i>	[m̪isra]	'boundary'

5.4. Diphthongs

5.4.1. NENA dialects

The original diphthongs **ay* and **aw* have contracted to /e/ and /o/ respectively in both Jewish and Christian NENA:

- J. Sanandaj *bela* (< **baytā*) 'house'; *mola* (< **mawtā*) 'death'
C. Sanandaj *besa, mosa*

5.4.2. Neo-Mandaic

Also in Neo-Mandaic the original diphthongs **ay* and **aw* have generally contracted (Häberl 2009: 61, 62, 74, 88–89). In unstressed open syllables the reflex of **ay* is /e/, e. g.

Khorramshahr

hemanuṭan [hema'nu:θæn] (< **haymānūtan*) 'our faith'

In stressed open syllables **ay* sometimes contracts to /i/, e. g.

Khorramshahr

inā ['i:nɔ] (< * 'aynā) 'eye'

In some words the process of contraction of this diphthong results in the hiatus sequence /ie/, which has developed by a breaking of the vowel **e* (i. e. **ay* > **e* > /ie/), e. g.

Khorramshahr

bieṭā ['biɛθɔ] (< **baytā*) 'house'

This breaking of the diphthong, which has been documented in recent years in both the Khorramshahr and the Ahvāz dialects, is not found in records of Neo-Mandaic from earlier in the twentieth century. This suggests that it may be a recent development under the influence of languages in contact, such as the Arabic dialect of Khuzestan, where the phenomenon is also found (Häberl 2009: 38–39, 74; Macuch 1993: 38). The diphthong **aw* generally contracts to /o/ or /u/ in open stressed syllables, e. g.

Khorramshahr	
<i>nodā</i> ['no:də] (< <i>nawdā</i>)	'quake'
<i>yumā</i> ['ju:mə] (< <i>yawmā</i>)	'day'

Diphthongs other than **ay* and **aw* and those resulting from inflectional processes or in loanwords are generally preserved in the Khorramshahr dialect, but they have a tendency to contract in the Ahvāz dialect (Häberl 2009: 51, 258, 2011: 728), e. g.

Khorramshahr		
Ahvāz	<i>gāw</i> [gɔv]	'in'

5.5. Stress

5.5.1. NENA dialects

The basic stress position in the Jewish cluster of dialects of the region is word-final, as in Jewish trans-Zab dialects elsewhere, whereas C. Sanandaj has basic penultimate stress:

J. Sanandaj	C. Sanandaj
<i>belá</i>	<i>bésa</i>
<i>qátəl</i>	<i>qátəl</i>

In the Jewish dialects, however, retraction of the word-final stress frequently occurs in various non-pausal contexts within an intonation group, e. g.

xá	'áxóna	<i>xet-àʃ'</i>
a	brother	other-3FS
'Another brother of hers' (Khan 2009: 53) ⁵		

In J. Sanandaj a short /ə/ or /u/ in an unstressed syllable may undergo devoicing. This renders the vowel /ə/ inaudible, e. g.

*šəmá > [ʃɪ'ma]	'heaven'	(Khan 2009: 40)
-----------------	----------	-----------------

⁵ The intonation group boundary is indicated by the symbol !. The nuclear stress is represented by a grave accent and non-nuclear stresses by acute accents.

When short /u/ is devoiced in this way, the lip-rounding gesture of the vowel remains, resulting in labialized pronunciation of the preceding consonant, e. g.

**duká* > [t^{hw'}k^ha] ‘place’ (Khan 2009: 43)

5.5.2. Neo-Mandaic

According to Häberl’s description of the Khorramshahr dialect (Häberl 2009: 77–81), in Neo-Mandaic the placement of stress depends on vowel quality and syllable structure. The stress generally falls on a vowel with the qualities [i], [u], or [ɔ], which Häberl (2009: 59) calls “tense vowels”, and preferably upon a closed word-final syllable. If the word-final syllable is not suitable, the stress moves towards the front of the word until it rests on a suitable syllable. Virtually all words have the stress on the final, penultimate or antepenultimate syllable.

The stress falls on the final syllable when it is closed and contains a tense vowel, e. g.

qəmahrəbāt [qə.mæh.re.'wɔ:t] ‘you destroy’

or when the word-final syllable does not contain a tense vowel but there are no suitable preceding syllables, e. g.

əkál [a.'χal] ‘he ate’

In words containing two syllables, if the final syllable is open or contains a lax vowel, the stress falls upon the penultimate syllable, if it contains a tense vowel or is closed, e. g.

<i>bábā</i>	[‘bɔ:.wɔ]	‘father’
<i>gábrā</i>	[‘gæv.rɔ]	‘man’

In words of three or more syllables, if neither the final nor the penultimate syllable is suitable for the stress placement, it recedes to the antepenultimate syllable, if this is closed or contains a tense vowel, e. g.

gatélnākon [ga.'t̪el.nɔ.χon] ‘I will kill you’

5.6. Syllable structure

5.6.1. NENA dialects

In C. Sanandaj consonant gemination has been lost, including after short unrounded /ə/, where it has been preserved in most other NENA dialects. When the gemination is lost after this vowel, the vowel remains short:

C. Sanandaj

<i>dáma</i> (< *dəm̥ma) ‘blood’	cf. C. Urmi <i>dəmma</i>
<i>lába</i> (< *ləbba) ‘heart’	cf. C. Urmi <i>ləbba</i>

A consonant that was geminated secondarily due to the assimilation of an adjacent consonant also loses its gemination in some cases. Such geminaton is, likewise, preserved in most NENA dialects:

C. Sanandaj

<i>šadráte</i> (< *šadrət̥te) ‘you send him’
<i>'átan</i> (< *'əttan < *'itlan) ‘we have’

Within the Jewish cluster of dialects of the region consonant gemination has been totally lost in J. Sanandaj, as in C. Sanandaj, but it has been retained after /ə/ in other Jewish dialects of the cluster in conformity with the majority of NENA dialects:

J. Sanandaj J. Săqəz

<i>dəmá</i>	<i>dəmmá</i>	‘blood’
<i>ləbá</i>	<i>ləbbá</i>	‘heart’

In some contexts a short /ə/ in an open syllable arising from loss of gemination is eliminated by resyllabification, e. g.

J. Sanandaj

<i>šəqlí</i> ‘I bought’	< *šqəli	< *šqəlli
<i>zábna</i> ‘I sell’	< *zābəna	< *zābənna

In J. Sanandaj, but not other Jewish dialects of the region, consonants that are geminated secondarily due to assimilation of another consonant also lose their gemination. In this case the vowel in the opened syllable is lengthened:

J. Sanandaj J. Săqəz

<i>garšéte</i>	<i>garšátte</i>	‘you pull him’ < *garšet-le
----------------	-----------------	-----------------------------

Some other features of syllable structure that are distinctive specifically of the western Iranian Jewish cluster of dialects include the following.

Within a word two vowels may follow one another without an intervening glottal stop //'. Such sequences should be analysed as diphthongs rather than two separate syllable nuclei. In word-final sequences of vowels the existence of the diphthong is reflected by the fact that stress that would normally be expected to be put on a final vowel nucleus (§5.5.1.) is retracted to the vowel preceding it, indicating that the final vowel is treated as non-syllabic, e. g.

J. Sanandaj

<i>huláe</i> [CV.CV̥] ‘Jews’
<i>'óa</i> [CVV] ‘that one’

In cases where stress is put on the second vowel of the sequence in word-final position, the diphthong should be interpreted as rising VV. This is found in sequences where the first of the two vowels is /o/. The non-syllabic status of /o/ is reflected by the fact that it is often realized as a semi-vowel [w], e. g.

J. Sanandaj

šoá [ʃwa] ‘seven’
n̥toá [ntwa] ‘high’

Sequences of three vowels with a medial /o/ occur in some forms. These should be interpreted as VVV, the medial vowel being the on-glide of a diphthong, the phonetic realization of which is often the semi-vowel [w], e. g.

J. Sanandaj *xǎoé* [xa'we] ‘to see’

Short /ə/ is often devoiced in unstressed syllables. This results in phonetic realizations such as the following in which the /ə/ vowel is inaudible:

J. Sanandaj *qətmá* [qʰətʰ'ma] ‘ash’
šəmšá [ʃəm'sa] ‘sun’

5.6.2. Neo-Mandaic

The gemination of consonants is widely attested in Neo-Mandaic. A distinction should be made between lexical gemination and phonological gemination, according to the terminology proposed by Häberl (2009: 75–77).

Lexical gemination includes geminated consonants inherited from earlier stages of the language, those formed through the total assimilation of one consonant to another, and gemination in loanwords, e. g.

<i>lébbā</i> ['lebbɔ]	‘heart’	< * <i>libbā</i>
<i>moqaddás</i> [moqad'das]	‘holy’	< Arabic <i>muqaddas</i>

Phonological gemination depends on vowel quality and stress. It occurs only after a stressed lax vowel, i. e. [e], [o] or [a]. The most common context where it is found is where an inflectional morpheme beginning with a vowel (typically a pronominal suffix) is added to a closed stressed syllable containing a lax vowel, e. g.

<i>həzon</i> [hə.'zon]	+ <i>a</i> [a]	> <i>həzónna</i> [hə.'zon.na]
see.PST.3PL	+ OBJ.3FS	see.PST.3PL.OBJ.3FS
‘They saw her’		

All consonants can undergo phonological gemination in this way, with the exception of the interdental fricative /θ/. In a context where the phonological gemination of /θ/ is expected, the cluster [χt] occurs instead of [θθ], e. g.

<i>eₜ</i> ['ɛθ]	+	<i>aₖ</i> [ax]	>	<i>éktak</i> ['ɛχ.tak]
COP		2MS		COP.2MS
'You are'				

6. Nominal morphology

6.1. Pronouns

6.1.1. NENA dialects

Both the Jewish cluster of dialects and C. Sanandaj have two series of demonstrative pronouns, which express near and far deixis respectively. Unlike many NENA dialects in other regions, they have not developed three series in which the original far deixis form is used specifically for anaphoric reference. The near and far deixis may be used with either deictic or with anaphoric function. The forms are provided in Table 1.

Table 1: Demonstratives in the Jewish and Christian NENA dialects

	J. Sanandaj	J. Bokan	J. Kerend	C. Sanandaj
Near deixis				
3ms.	'ay, 'éa, 'éxa	'ay, 'áya	'ay, 'áya	'ay, 'áya, 'íya, 'ayánān, 'iyánān
3fs.	'ay, 'éa, 'éxa	'ay, 'áya	'ay, 'áya	'ay, 'áya, 'íya, 'ayánān, 'iyánān
3pl.	'ayní, 'onyé, 'onyexáe	'onye	'onyé	'ān, 'áni, 'anínān
Far deixis				
3ms.	'o, 'óa, 'óxa	'o, 'óya	'o, 'áwa	'o, 'áwa, 'awánān
3fs.	'o, 'óa, 'óxa	'o, 'óya	'o, 'áwa	'o, 'óya, 'oyánān
3pl.	'óni, 'onyé, 'onyexáe	'óni, 'onyé	'óni, 'onyé	'ōn, 'óni, 'onínān

As can be seen, most of the demonstrative pronouns in the dialects have short and long forms. The long forms tend to be used only when the pronoun is independent and not combined with a noun. Only singular forms can be used in combination with a noun. These are the shortest forms in the list presented above. These singular adnominal forms are used also with plural nouns, e. g.

J. Sanandaj

- | | |
|-----------------|----------------|
| <i>'ay gora</i> | 'this man' |
| <i>'ay naše</i> | 'these people' |

- | | |
|----------------|----------------|
| <i>'o gora</i> | 'that man' |
| <i>'o naše</i> | 'these people' |

A few minor differences are found across the Jewish dialects of the region with regard to the contraction of diphthongs in the pronouns (J. Sananadaj *'ea* < **'ayha*, *'óa* < **'awha* vs. J. Săqəz, J. Kerend *'áya* < **'ayha*, *'áwa* < **'awha*). In both the Jewish dialects and C. Sanandaj there is a levelling of gender in the singular. In C. Sanandaj, however, differences in contraction of a diphthong are exploited to mark a distinction in gender in the far deixis form (*'áwa* ms. < **'awha*, *'óya* fs. < **'awha*). The 1st and 2nd person pronouns are given in Table 2.

Table 2: First and second person pronouns in Jewish and Christian Sanandaj

	J. Sanandaj	C. Sanandaj
2ms.	<i>'āt</i>	<i>'ayət</i>
2fs.	<i>'āt</i>	<i>'ayat</i>
2pl.	<i>'axtu</i>	<i>'axtoxən</i>
1s	<i>'ana</i>	<i>'ana</i>
1pl.	<i>'axni</i>	<i>'axni</i>

C. Sanandaj expresses gender distinction in the 2nd singular with innovative forms, found widely elsewhere in NENA, formed by the attachment of suffixes taken from the paradigm of D-suffixes that are used in verbal inflection.

6.1.2. Neo-Mandaic

Neo-Mandaic has two series of demonstrative pronouns in the singular, expressing near and far deixis (Häberl 2009: 161–163). There is only one plural demonstrative pronoun, which is historically a far deixis form, but now is neutral as to distance. The singular pronouns are neutral as to gender and occur in an independent and an adnominal dependent form, which are shown in Table 3.

Table 3: Neo-Mandaic demonstratives

		Independent	Adnominal
Near deixis			
3s.	<i>āhā</i>	<i>ā</i>	‘this’
3pl.	<i>ahni</i>	—	‘these’
Far deixis			
3s	<i>aku</i>	<i>ak</i>	‘that’
		<i>tā</i>	
3pl.	<i>ahni</i>	—	‘those’

The use of the singular adnominal demonstratives is extended to plural nouns, since the plural demonstrative does not have an adnominal form, e. g.

- ā bietā* ‘this house’
ā šeršānā ‘these religions’

In addition to the near and far deixis demonstratives, there is another series of 3rd person pronouns that functions as anaphoric personal pronouns. This differs from the NENA dialects of the region, in which the deictic pronouns have taken over the function of the anaphoric pronoun. The full paradigm of independent personal pronouns is as follows. Some pronouns have two forms, which Häberl (2009: 156) terms “lexical” and “contextual” respectively (Table 4).

Table 4: Lexical and contextual pronouns in Neo-Mandaic

	Lexical	Contextual
3ms.	<i>huwi</i>	<i>huw</i>
3fs.	<i>hidā</i>	<i>hid</i>
3pl.	<i>honni</i>	—
2s.	<i>āt</i>	—
2pl.	<i>atton</i>	—
1s.	<i>anā</i>	<i>an</i>
1pl.	<i>ani</i>	<i>an</i>

The contextual forms, the final vowel of which has been elided, are subject pronouns before a verb that do not carry the nuclear stress of the utterance (a). The lexical form of the pronoun is typically used when it takes the nuclear stress of the utterance and has a contrastive function (b):

- (a) *an yəhm-it pahli-d-ak*
 1SG sit.PST-1SG beside-GEN-2SG.M
 'I sat beside you.'
- (b) *anā yəhm-it pahli-d-ak*
 1SG sit.PST-1SG beside-GEN-2SG.M
 'I (was the one who) sat beside you.'

6.2. Pronominal suffixes

6.2.1. NENA dialects

The pronominal suffixes attached to nouns and prepositions are shown in Table 5.

Table 5: Pronominal suffixes in Jewish and Christian Sanandaj

	J. Sanandaj	C. Sanandaj
3ms.	- <i>ef</i>	- <i>e</i>
3fs.	- <i>af</i>	- <i>e</i>
3pl.	- <i>u</i> , - <i>un</i>	- <i>u</i>
2ms.	- <i>ox</i>	- <i>ox</i>
2fs.	- <i>ax</i>	- <i>ax</i>
2pl.	- <i>ăxun</i>	- <i>oxən</i>
1s	- <i>i</i>	- <i>i</i>
1pl.	- <i>an</i> , - <i>ăni</i>	- <i>an</i>

The 3rd person singular suffixes in the Jewish dialects have an augment *-f*, which derives historically from *-u*; cf. the forms *-eu* 3ms. and *-aw*, which occur in several NENA dialects in other regions (chapter 3.4, §6.1). In C. Sanandaj the original 3ms form *-e* has been extended to refer also the 3fs.

The variant form of the 3pl. suffix in J. Sanandaj *-un*, which preserves the original final */n/* of the pronoun (< *-hun), is used only before the co-ordinative enclitic *-u*, e. g. *xa-dana mənun-u* 'one of them and ...' (Khan 2009: 61).

6.2.2. Neo-Mandaic

The pronominal suffixes attached to nouns and prepositions and also to verbs, when expressing the object, are given in Table 6.

Table 6: Pronominal suffixes in Neo-Mandaic

3ms.	<i>-i</i>
3fs.	<i>-a</i>
3pl.	<i>-u</i>
2ms.	<i>-ak</i>
2fs.	<i>-ek</i>
2pl.	<i>-kon</i>
1s.	<i>-e</i>
1pl.	<i>-an</i>

When attached to nouns these function as genitive suffixes, e. g.

emm-e
mother-1SG
'my mother'

The genitive pronominal suffixes are attached to loanwords with an intervening genitive particle *d-* (Häberl 2009: 158), e. g.

'umər-d-e
age-GEN-1SG
'my age'

6.3. Inflection of nouns and adjectives

6.3.1. NENA dialects

The inflection of nouns and adjectives in the NENA dialects of the region follow the same basic patterns as other NENA dialects (chapter 3.4, §6.3).

In the Jewish dialects the feminine singular ending has the form *-da* or *-la* in words in which the original ending was **-ta* with an interdental fricative. The ending *-da* occurs in a few nouns whose stem ends in one of the sonorant consonants /l/, /n/ or /r/:

J. Sanandaj *qarda* (cf. C. Barwar *qarθa*) 'cold'
šənda (cf. C. Barwar *šənθa*) 'sleep'

The ending *-la* occurs in nouns the stem of which ends in a vowel or the sonorant consonants /w/ and /m/:

J. Sanandaj	<i>šala</i> (cf. C. Barwar <i>šaθa</i>) ‘fever’
	<i>šwila</i> (cf. C. Barwar <i>šwiθa</i>) ‘bed’
	<i>xmala</i> (cf. C. Barwar <i>xmaθa</i>) ‘mother-in-law’

In C. Sanandaj the feminine ending has the form *-sa* in these contexts: *šənsa* ‘sleep’, *šasa* ‘fever’, *šwisa* ‘bed’, *xmasa* ‘mother-in-law’.

In both the Jewish dialects and C. Sanandaj the common nominal plural ending *-e* has been analogically extended to the feminine plural ending **-ātā*. In the Jewish dialects this results in the form *-ale* or *-ae* and in C. Sanandaj the resultant form is *-ase*:

J. Sanandaj	C. Sanandaj		
<i>'ilale</i>	<i>'idase</i>	‘hands’	< * <i>'idātā</i>
<i>šukyale</i>	<i>škyase</i>	‘testicles’	< * <i>šekyātā</i>
<i>belāwāe</i>	<i>baswase</i>	‘houses’	< * <i>baytāwātā</i>
<i>malāwāe</i>	<i>maswase</i>	‘villages’	< * <i>mātāwātā</i>

Both the Jewish dialects and C. Sanandaj have borrowed a definite article suffix from Sorani Kurdish with the form *-āke*, e. g. J. Sanandaj *bela* ‘house’, *belāke* ‘the house’; C. Sanandaj *besa* ‘house’, *besāke* ‘the house’. The form *-āke* corresponds to the oblique form of the Kurdish definite article (*-ākay*) rather than the nominative form (*-āka*).

In all NENA dialects indefiniteness is expressed by the cardinal numeral ‘one’ (*xa*), e. g. J. Sanandaj *xa bela* ‘a house’.

6.3.2. Neo-Mandaic

In Neo-Mandaic the original Aramaic nominal inflection *-ā*, is retained, but the various Aramaic plural morphemes have been mostly replaced by the morpheme *-ān-*, which may be identified with the Persian plural morpheme of the same form (Häberl 2011: 730). This plural morpheme is placed before the nominal inflection *-ā*. It is used with nouns of both masculine and feminine gender, e. g.

Gender	Singular	Plural
m.	<i>kədāb-ā</i>	‘book’
f.	<i>id-ā</i>	‘hand’

The Aramaic plural morpheme *-āt-* (sometimes augmented by the glides /w/ or /y/) has been preserved in some nouns. It regularly occurs in feminine nouns that have retained the Aramaic feminine singular morpheme *-t-* in the singular, e. g.

Gender	Singular	Plural
f.	<i>tur-t-ā</i>	‘cow’

It is occasionally found in other nouns of feminine gender without the feminine morpheme *-t-* in the singular, e. g.

f.	<i>biet-ā</i>	'house'	<i>biet-wāt-ā</i>	'houses'
f.	<i>hewān-ā</i>	'animal'	<i>hewān-yāt-ā</i>	'animals'

Most loanwords are inflected with the Persian plural suffix *-(h)āž*, e. g.

Gender	Singular	Plural
m.	<i>jihel</i>	'child'

Some frequent nouns have irregular plurals that have a different stem from the singular and in some cases do not contain a plural morpheme (Häberl 2009: 130), e. g.

Gender	Singular	Plural
m.	<i>gabr-ā</i> 'man'	<i>gobr-ā ~ gobr-ān-ā</i> 'men'
f.	<i>ett-ā</i> 'woman'	<i>enš-ā ~ enš-ān-ā</i> 'women'
f.	<i>hāt-ā</i> 'sister'	<i>ah-wāt-ā</i> 'sisters'

Neo-Mandaic has no definite article but it uses various strategies borrowed from Persian to mark indefiniteness. A referentially indefinite noun, which has a specific or non-specific referent, may be marked by the suffix *-i*, by the preposed particle *ya* (< Persian *yak* 'one') or by both (Häberl 2009: 144). The suffix *-i* replaces the nominal inflection *-ā*, e. g.

<i>barnāš-ā</i>	'the person'	<i>barnāš-i</i>	'a person'
<i>barnāš-ān-ā</i>	'the people'	<i>barnāš-ān-i</i>	'some people'
<i>mend-ā</i>	'the thing'	<i>mend-i</i>	'a thing, something'
<i>ya kədāb</i>	<i>mandāyī</i>		
a book	Mandaic		
	'a Mandaic book'		
<i>ya tājer-i</i>			
a merchant	INDEF		
	'a merchant'		

This indefinite marker is not used productively in NENA, but the word *māndi* 'thing, something' is a common word in all NENA dialects. It is possible that this spread into NENA by areal diffusion from Mandaic at an earlier period.

Plural markers are used in principle only with referential nouns. They do not occur on nouns denoting a generic class, e. g.

<i>barnāš-ā</i>	'the person' (singular referential)/'people' (generic)
-----------------	--

6.4. Nominal annexation

6.4.1. NENA dialects

Nominal annexation involves the formation of a complex nominal phrase consisting of a head noun and a dependent noun. In C. Sanandaj the Aramaic genitive particle *d* is suffixed to the head noun, which generally results in the ending *-əd*:

- | | |
|-------------|--|
| C. Sanandaj | <i>'oṭaqəd qaša</i> ‘the room of the priest’ |
| | <i>sa'atəd mose</i> ‘the hour of his death’ |
| | <i>našəd 'asra</i> ‘the people of the town’ |

When a demonstrative pronoun is in the dependent phrase, both the demonstrative and the head noun have *d*. The form of the demonstrative beginning with *d* can be regarded as an oblique allomorph of the pronoun:

- | | |
|-----------------|----------------------------|
| šəməd do marhūm | ‘the name of the deceased’ |
|-----------------|----------------------------|

In the Jewish dialects the most common way of annexing one nominal to another in a genitive relationship is simply to juxtapose the two. The original Aramaic genitive particle *d-*, which regularly occurs in such constructions in other NENA dialects, is rarely used in the Jewish dialects of this region:

- | | |
|-------------|---|
| J. Sanandaj | <i>bela bäruxi</i> ‘the house of my friend’ |
|-------------|---|

The *d-* particle has survived only in demonstrative pronouns. The form of the pronoun beginning with *d-* can be regarded as an oblique allomorph of the pronoun:

- | | |
|-------------|---|
| J. Sanandaj | <i>bela do naša</i> ‘the house of that man’ |
|-------------|---|

In the Jewish dialects in the south of the region, such as J. Kerend, the *d-* is generally omitted even on demonstrative pronouns:

- | | |
|-----------|---|
| J. Kerend | <i>bela 'o naša</i> ‘the house of that man’ |
|-----------|---|

Occasionally the Iranian *ezafe* particle *-e* is suffixed to the head noun in the Jewish dialects:

- | | |
|-------------|---|
| J. Sanandaj | <i>belá-e bäruxi</i> ‘the house of my friend’ |
| | <i>belá-e do naša</i> ‘the house of that man’ |

In a few closely-knit phrases in both Jewish and Christian NENA the head noun of an annexation phrase has no inflectional ending, e. g.

- | | |
|-------------|--|
| J. Sanandaj | <i>reš-šata</i> ‘beginning of the year (= New Year)’ |
| C. Sanandaj | <i>mār-kalba</i> ‘the master of the dog’ |

6.4.2. Neo-Mandaic

In Neo-Mandaic nominal annexation is expressed by the elision of the nominal ending *-ā*. This occurs when a noun is immediately followed by another noun or adjective expressing a genitive or attributive relationship (Häberl 2009: 91, 132, 2011: 730), e. g.

<i>bietā</i>	‘house’	<i>biet bābe</i>	‘the house of my father’
<i>bietā</i>	‘house’	<i>biet ba ‘id</i>	‘the distant house’

If a phrase containing an adjective is indefinite, the indefinite suffix *-i* may be attached to the adjective at the end of the phrase, e. g.

<i>gazgān honin-i</i>	‘a small cooking pot’ (Häberl 2009: 147)
-----------------------	--

The distribution of such shortened noun forms replicates the pattern of distribution of Persian *ezafe* constructions. The Persian *ezafe* particle itself may be used, e. g.

<i>kol</i>	<i>dukkān=e</i>	<i>Amrik-ān</i>
all	shop=EZ	American-PL
‘all the American shops’		

This is sporadically found in the Khorramshahr dialect of Neo-Mandaic (Häberl 2009: 133–134) but is more frequent in the Ahvāz dialect (Macuch 1993: 60–61). The shortened ‘contextual’ forms, to use Häberl’s terminology, are analogous to the shortened ‘contextual’ forms of pronouns (§6.1.2.).

A genitive relationship between nouns may also be expressed by the particle *d-*. In such cases the head noun is in its lexical form with the ending *-ā*, e. g.

<i>qanāyā</i>	<i>d=kaspā</i>
smith	of=silver
‘silver smith’ (Häberl 2009, 231)	

7. Verbal morphology

7.1. NENA dialects

The basic principles of verbal derivational morphology and inflection in the NENA dialects of the region are the same as those that have been described for the NENA dialects of Iraq (chapter 3.4, §7). Here we shall concentrate on some selected features.

In C. Sanandaj the original NENA system of three derivative stems has undergone some degree of levelling, with the result that stem II is partially assimilated to the other stems. The present base of stem II is identical in form to that of stem

I, whereas the past base has the form of that of stem III. The only base that is distinctive to stem II is the imperative.

As in other NENA dialects (chapter 3.4, §7), the present and past inflectional bases are inflected for person and number by two sets of suffixes, termed D-suffixes (i. e. direct suffixes) and L-suffixes (i. e. oblique suffixes deriving historically from prepositional phrases containing the dative/agentive preposition *l-*). These suffixes indicate the grammatical relations of verbal arguments in the clause. In C. Sanandaj D-suffixes mark the subject of the present base and L-suffixes the subject of the past base, as shown in Table 7.

Table 7: Verb stems with D- and L-suffixes in Christian Sanandaj

Stem I <i>g-r-š</i> I ‘to pull’	Present base: <i>garəš</i> -D Past base: <i>grəš</i> -L Imperative: <i>groš</i>
Stem II <i>š-d-r</i> II ‘to send’	Present base: <i>šadər</i> -D Past base: <i>məšdər</i> -L Imperative: <i>šadər</i>
Stem III <i>x-l-p</i> III ‘to exchange’	Present base: <i>maxləp</i> -D Past base: <i>məxлəp</i> -L Imperative: <i>maxləp</i>

The morphology of the verbal derivative stems of the Jewish cluster of dialects has undergone considerable development. The main surviving stems of the original three-stem system of NENA are Stem-I (derived from the *pe’al* of earlier Aramaic) and Stem-III (derived from the causative *’ap’el* of earlier Aramaic). As in C. Sanandaj, stem II has virtually completely merged in form with Stem-I, but a distinction in the imperative form justifies identifying it as a separate stem. Another development has taken place that is not found elsewhere in NENA, with the exception of the closely related dialect of J. Sulemaniyya. In each stem the vocalic pattern of the bases differs according to whether the verb is used transitively or intransitively. Broadly speaking the original vocalic patterns of the bases of the causative Stem-III have been extended to transitive Stem-I verbs. Conversely the intransitive Stem-I verbs preserves the original vocalic patterns of Stem-I and this has been extended to intransitive verbs in Stem-III. Similar vocalic patterns are found in the transitive and intransitive verbs of Stem-II. The subject of all present bases is expressed by D-suffixes. There is, however, a split in the inflection of past

bases. Transitive past bases are inflected by L-suffixes but intransitive past bases are inflected by D-suffixes, as shown in Table 8, while Table 9 shows the forms of the suffixes themselves.

Table 8: Verb stems with D- and L-suffixes in Jewish Sanandaj

Stem I		
Present base:	<i>g-r-š</i> I ‘to pull’ (transitive)	<i>s-m-x</i> I ‘to stand’ (intransitive)
Past base:	<i>gärəš</i> -D	<i>sāməx</i> -D
Resultative participle:	<i>grəš</i> -L	<i>smix</i> -D
Imperative:	<i>gärəš</i> , <i>gruš</i>	<i>smux</i>
Infinitive:	<i>gärōše</i>	<i>sāmoxe</i>

Stem II		
<i>z-b-n</i> II ‘to sell’		
	Transitive (active)	Intransitive (anticausative)
Present base:	<i>zábən</i> -D	—
Past base:	<i>zbən</i> -L	<i>zbin</i> -D
Resultative participle:	<i>zəbna</i>	<i>zbina</i>
Imperative:	<i>zábən</i>	—
Infinitive:	<i>zăbone</i>	—

Stem III		
	<i>m-rš-x</i> ‘to cause to walk’ (transitive)	<i>m-sk-r</i> ‘to become lost’ (intransitive)
Present base:	<i>marxəš</i> -D	<i>maskər</i> -D
Past base:	<i>mərxəš</i> -L	<i>məskir</i> -D
Resultative participle:	<i>məršxa</i>	<i>məskira</i>
Imperative:	<i>márxəš</i>	<i>móskur</i> —
Infinitive:	<i>marxoše</i>	<i>maskore</i> —

Table 9: D- and L-suffixes in Jewish Sanandaj

	D-suffixes	L-suffixes
3ms	-Ø	-le
3fs	-a	-la
3pl	-i	-lu, -lun
2ms	-et	-lox
2fs	-at	-lax
2pl	-etun	-läxun
1ms	-na	-li
1fs	-an	-li
1pl	-ex	-lan

It is a common feature of the NENA dialects that the subject of present base verbs is expressed by D-suffixes and the pronominal object by oblique L-suffixes. The Jewish NENA dialects of western Iran, however, deviate from this pattern in one detail, in that the pronominal object of 1st person singular present base verbs are expressed by suffixes from the series that are attached to nouns to express possessors (see §6.2.1), e. g.

J. Sanandaj (Khan 2009: 154–155)

gārəš-Ø-le

pull.PRS-D.3MS-L.3MS

‘he pulls him’

garš-a-le

pul.PRS-D.3FS-L.3MS

‘she pulls him’

garš-i-le

pull.PRS-D.3PL-L.3MS

‘they pull him’

gārəš-n-ef

pull.PRS-D.1MS-POSS.3MS

‘I (ms) pull him’

garš-ăn-ef

pull.PRS-D.1FS-POSS.3MS

‘I (fs) pull him’

In the Jewish dialects of the region the D-suffixes are used also to mark the object of transitive verbs, at least when the object is 3rd person. The alignment of clauses with verbs constructed on past bases in these dialects, therefore, is ergative. If the subject or object are specified by noun phrases in the clause, these are cross-referenced by the verbal suffixes. This is illustrated in the following examples, which show how the subject of an intransitive verb (1a) aligns with the object of a transitive verb (1b). In each case the common marker is the 3fs D-suffix on the verb:

- (1) J. Sanandaj

a. *brat-i smix-a.*

daughter-my stand.PST-D.3FS

‘My daughter (S) stood up.’

b. *baruxăwal-i brat-i gərš-a-lu.*

friends-my daughter-my pull.PST-D.3FS-L.3PL

‘My friends (A) pulled my daughter (O).’

Unergative intransitive verbs, with agentive subjects, take L-suffixes on the past base. The dialects, therefore, should be characterized as Split-S, according to the

terminology of Dixon (1994: 71), rather than canonically ergative. This applies, for example, to verbs expressing the emission of sounds:

- (2) J. Sanandaj
kalba nwəx-le.
 dog bark.PST-L.3MS
 ‘The dog barked.’

Intransitive verbs with D-suffixes may be termed unaccusative since they undergo a change of state and are typically not agentive. They include verbs of movement, which express a change in locational state rather than change in configurational state, e. g.

- (3) J. Sanandaj
brat-i zil-a.
 daughter-my go.PST-D.3FS
 ‘My daughter went.’

The division between unergative and unaccusative in verbs of emission of sound is lexicalized slightly differently across the various dialects of Jewish NENA of the region, e. g.

	Unergative (L-suffix)	Unaccusative (D-suffix)
‘She coughed’	J. Sanandaj (<i>šəh-la</i>) J. Tikab (<i>šəl-la</i>) J. Kerend (<i>šhəl-la</i>)	J. Qar Hasan (<i>šhil-a</i>) J. Bokan (<i>šhil-a</i>)
‘He sneezed’	J. Sanandaj (<i>təp-la</i>) J. Tikab (<i>tpəl-la</i>) J. Kerend (<i>tpəl-la</i>)	J. Bokan (<i>tpil-a</i>) J. Qar Hasan (<i>tpil-a</i>)

In the Jewish dialects the D-suffixes on the past base of transitive verbs can express only the 3rd person object. A 1st or 2nd person object of a past base verb must be expressed by an independent object phrase outside of the verbal complex:

- (4) J. Sanandaj
 a. *grəš-le* *'əl-i.*
 pull.PST-L.3MS OBL-1S
 ‘He pulled me.’
 b. *grəš-le* *'əl-ox.*
 pull.PST-L.3MS OBL-2MS
 ‘He pulled you.’

In C. Sanandaj past base verb forms cannot express any pronominal objects by D-suffixes. Another construction must be used consisting of a past converter prefix *təm-* and the present base, which expresses the pronominal object by L-suffixes:

(5) C. Sanandaj

- a. *təm-garš-i-le.*
PST-pull.PRS-D.3PL-L.3MS
'They pulled him.'
- b. *təm-garš-a-le.*
PST-pull.PRS-D.3FS-L.3MS
'She pulled him.'

In both the Jewish dialects and C. Sanandaj ditransitive verbs cannot take two pronominal object suffixes. In cases where there is a direct and an indirect pronominal object, the indirect one must be expressed by an independent prepositional phrase:

(6) J. Sanandaj

- | | |
|--------------------|----------------|
| <i>kwi-le</i> | <i>'əl-ef.</i> |
| give.PRS.3PL-L.3MS | to-3MS |
- 'They give it (m.) to him.'

(7) C. Sanandaj

- | | |
|--------------------|----------------|
| <i>kewi-le</i> | <i>tlas-e.</i> |
| give.PRS.3PL-L.3MS | to-3MS |
- 'They give it (m.) to him.'

Calques of phrasal verbs in the languages in contact containing a light verb and noun are common in both the Jewish and Christian dialects, e. g.

(8) J. Sanandaj (Khan 2009: 153)

- hāz 'ol* 'he desires' (literally: he makes desire; cf. Kurdish *haz kirdin*)
komak 'ol 'he helps' (literally: he does help; cf. Persian *komak kardan*).

The inflection of the enclitic copula in the Jewish dialects of the region corresponds to that of the irrealis verb *h-w-y* 'to be' throughout the paradigm. This represents a more advanced state of assimilation of the enclitic copula to verbal inflection than is found in other NENA dialects. C. Sanandaj exhibits a paradigm of the enclitic copula that is more usual in NENA whereby the 1st and 2nd person forms of the copula correspond to verbal endings, but the 3rd person forms have not undergone levelling and they remain distinct in form from verbal inflection:

Table 10: Enclitic copulas in Jewish and Christian Sanandaj

J. Sanandaj		C. Sanandaj		
	Enclitic Copula	Verb <i>h-w-y</i>	Enclitic Copula	Verb <i>h-w-y</i>
3ms.	=ye	<i>hăwe</i>	=ile	<i>hawe</i>
3fs.	=ya	<i>hawya</i>	=ila	<i>hoya</i>
3pl.	=yen	<i>hăwen</i>	=ilu	<i>hawi</i>
2ms.	=yet	<i>hăwet</i>	=yet	<i>hawet</i>
2fs.	=yat	<i>hawyat</i>	=yat	<i>hoyat</i>
2pl.	=yetun	<i>hăwetun</i>	=iton	<i>hawiton</i>
1ms.	=yena	<i>hăwena</i>	=yen	<i>hawen</i>
1fs.	=yan	<i>hawyan</i>	=yan	<i>hoyan</i>
1pl.	=yex	<i>hăwéx</i>	=yox	<i>hawox</i>

Most of the Split-S Jewish dialects of the region express perfects by combining the resultative participle with the copula. The morphological form of the participle differs according to whether the verb is transitive or intransitive unaccusative. This distinction in the form of participle is found in the J. Sulemaniyya dialect of eastern Iraq but not in the vast majority of NENA dialects that form resultative perfects with a participle and copula construction (chapters 2.5, §7. and 3.4, §7). The cluster of Jewish dialects in western Iran differ from all other NENA dialects, including J. Sulemaniyya, in that the participle and the copula in transitive clauses agree with the object and not with the subject (Khan 2009: 90–92, 323–326), e. g.

(9) J. Sanandaj

- a. *baxt-ăke qimta=ya*
woman-the rise.PTCP.FS-COP.3FS
'The woman has risen.'
- b. *'o-gora baxt-ăke grəšta=ya*
that-man woman-the pull.PTCP.FS-COP.3FS
'The man has pulled the woman.'

It can be seen that the transitive subject argument in the transitive construction in (9b) has no explicit agreement marking on the verb in the form of an ergative L-suffix. When there is no transitive subject argument nominal in the clause, the subject can be interpreted as 3rd person of any gender or number:

(10) J. Sanandaj

- baxt-ăke grəšta=ya*
woman-the pull.PTCP.FS-COP.3FS
'He/she/they has/have pulled the woman.'

Such transitive perfect constructions with zero subject can be used only when the subject is 3rd person. By contrast, the perfect construction with a copula is available for subjects of all persons in intransitive unaccusative clauses:

J. Sanandaj

3ms	<i>smixá=ye</i>	'He has stood up'
3fs	<i>smixtá=ya</i>	'She has stood up'
2ms	<i>smixá=yet</i>	'You (ms.) have stood up'
2fs	<i>smixtá=yat</i>	'You (fs.) have stood up'
1ms	<i>smixá=yena</i>	'I (ms.) have stood up'
1fs	<i>smixtá=yan</i>	'I (fs.) have stood up'

An irrealis perfect may be formed by replacing the copula by the paradigm of the irrealis form of the verb *h-w-y* 'to be'. In the irrealis perfect construction in transitive clauses the participle and copula agree with the object, as in the indicative perfect, but, unlike in the indicative perfect, the subject is marked by an ergative L-suffix. Since there is explicit subject marking, the construction can be used with subjects of all persons (Khan 2009: 92–94):

(11) J. Sanandaj

- a. *'o-gora baxt-äke grəšta-hawya-le*
that-man woman-the pull.PTCP.FS-be.IRREALIS.3FS-L.3MS
'That man may have pulled the woman.'
- b. *'āt baxt-äke grəšta-hawya-lox*
you (ms.) woman-the pull.PTCP.FS-be.IRREALIS.3FS-L.2MS
'You (ms.) may have pulled the woman'
- c. *'ana baxt-äke grəšta-hawya-li*
I woman-the pull.PTCP.FS-be.IRREALIS.3FS-L.1S
'I may have pulled the woman'

In C. Sanandaj the perfect is not expressed by a resultative participle but rather by a construction in which the particle *gi-* is prefixed to a past base verb, e. g.

C. Sanandaj

<i>grəš-le</i>	<i>gi-grəš-le</i>
'he pulled'	'he has pulled'

This particle *gi-* is likely to be a reduced form of the copula with an indicative prefix < **k-i*.

Perfектs denote the resultative and also can have an evidential function used to express reported events or events in the remote past (Khan 2009: 296, 2012).

Neither Jewish nor Christian NENA have prefixed particles to mark future tense. In Jewish NENA the basic present base serves to cover general present, progressive present and future:

J. Sanandaj *gărəš* ‘he pulls/he is pulling/he will pull’

Both Jewish and Christian NENA have vestiges of an indicative prefix *k-* (or *g-*) which is attached to the present base of a few verbs with weak initial radicals when it expresses realis:

	Present realis	Present irrealis
C. Sanandaj	<i>k-azəl</i>	<i>'azəl</i>
J. Sanandaj	<i>g-ezəl</i>	<i>hezəl</i>

In J. Sanandaj the present progressive may be expressed by combining an inflected present base form of the verb with a preceding infinitive form (Khan 2009: 332), e.g.

J. Sanandaj
găroše *gărəš*
 pull.INF pull.PRS.3MS
 ‘he is pulling’

In constructions with verbs that take a *k-/g-* prefix in the present indicative, the *k-* is reduplicated on the infinitive, e.g.

J. Sanandaj
g-zala *g-ezəl*
 IND-go.INF IND-go.PRS.3MS
 ‘he is going’

In C. Sanandaj the progressive is expressed by combining an inflected form of the copula clitic with a present base verb (Panoussi 1990: 118), e.g.

gareš=ile
 pull.PRS.3MS=COP.3MS
 ‘he is pulling’

garš-en=yen
 pull.PRS-D.1MS=COP.1MS
 ‘I am pulling’

Some Jewish NENA dialects express the progressive by combining the present indicative with a possessive expression (*hit-e* ‘he has’, *hit-a* ‘she has’ etc.), e.g.

J. Kerend (Hopkins 1999: 325)
hit-e *k-e*
 POSS-3MS IND-come.PRS.3MS
 ‘he is coming’

7.2. Neo-Mandaic

Neo-Mandaic has preserved the suffix conjugation of earlier Aramaic to express the perfective past. The historical active participle expresses both the imperfective realis and irrealis. The realis is distinguished from the irrealis by a prefixed indicative particle, which has the form *qə-* in Khorramshahr (Häberl 2009: 180) and *qa-* in Ahvāz (Macuch 1993: 69). This is analogous, both in form and function, to the indicative particles *k-* and *g-* of the NENA dialects, though it may be of a different etymology. The general profile of the Neo-Mandaic and NENA verbal systems in relation to Middle Aramaic, the term generally used to refer to the historical stage of Aramaic before that of the modern dialects, is as follows (shading indicates continuity of Middle Aramaic forms in the Neo-Aramaic dialects):

	Middle Aramaic	Neo-Mandaic	NENA
Perfective past	<i>qtal</i>	<i>gətal</i>	<i>q̄til-le</i>
Imperfective realis	<i>qātel</i>	<i>qə-gātel</i>	<i>k-qātel</i>
Imperfective irrealis	<i>yiqtul</i>	<i>gātel</i>	<i>qaṭel</i>
Imperative	<i>qtul</i>	<i>gətol</i>	<i>qtol</i>

Neo-Mandaic has lost the Middle Aramaic imperfective irrealis form (*yiqtul*), but its verbal system is still more archaic than that of the NENA dialects, which have lost both the Middle Aramaic imperfective irrealis (*yiqtul*) and the perfective past (*qtal*).

The verbal stems include the historical *pe'el* (also known as the G-stem), *pa'el* (agentive or denominational D-stem), *'ap'el* (causative C-stem) forms as well as the T-stems (expressing middle and passive voices) *'etpe'el* (tG stem), *'etpa'al* (tD stem) and *'ettap'al* (tC stem). This is a further aspect of the verbal system that is more conservative than NENA, in which the T-stems have been lost. The prefixed *t* morpheme of the historical T-stems has, however, been assimilated in all verbal roots except those that have a sibilant as initial radical, with which it has been metathesized, e. g. *estəbā* ‘to be baptized’ (<*s-b-w/y*>). An overview of the basic forms (3ms perfective and imperfective, ms. imperative) across the various derivative stems is provided in Table 11 (based on Häberl 2011: 733).

Table 11: Verbal stems in Neo-Mandaic

Stem	Perfective	Imperative	Imperfective	
G-stem	<i>gətal</i>	<i>gətol</i>	<i>gātel</i>	‘to kill’
tG-stem	<i>epseq</i>	<i>epseq</i>	<i>mepseq</i>	‘to be cut off’
D-stem	<i>kammer</i>	<i>kammer</i>	<i>əmkammer</i>	‘to return’
tD-stem	<i>kammar</i>	<i>kammar</i>	<i>mekammar</i>	‘to turn back’
C-stem	<i>ahreb</i>	<i>ahreb</i>	<i>mahreb</i>	‘to destroy’
tC-stem	<i>ettar</i>	<i>ettar</i>	<i>mettar</i>	‘to wake up’

Compound verbal constructions consisting of light verbs expressing person, tense, mood and aspect combined with a preceding noun or adjective expressing lexical content, are very frequent. These are mostly calques on Persian light verb constructions. The most common light verbs are *əbad* ‘to do’, *əhab* ‘to give’, *məhā* ‘to hit’ and *tammā* ‘to become’ (Häberl 2009: 227), e. g.

- hitrā əbad* ‘to celebrate’
tamām tammā ‘to be finished’

The Neo-Mandaic perfective and imperfective forms are inflected with suffixes as follows in the Khorramshahr dialect, as shown in Table 12.

Table 12: Verbal inflection in Neo-Mandaic

	Perfective	Imperfective
3ms	<i>gəṭal-</i> Ø	<i>qə-ǵaṭel</i>
3fs	<i>gəṭl-at</i>	<i>qə-ǵaṭl-ā</i>
3pl	<i>gəṭal-yon</i>	<i>qə-ǵaṭl-en</i>
2s	<i>gəṭal-t</i>	<i>qə-ǵaṭl-et</i>
2pl	<i>gəṭal-ton</i>	<i>qə-ǵaṭl-etton</i>
1s	<i>gəṭl-it</i>	<i>qə-ǵaṭel-nā</i>
1pl	<i>gəṭal-ni</i>	<i>qə-ǵaṭl-enni</i>

Macuch (1993: 68) reports the occurrence of distinct suffixes of the perfective for the 2fs, 3fpl and 2fpl in the Ahvāz dialect, but these do not occur in the material gathered by Häberl (2009: 180), who gives the paradigm above.

Pronominal objects on perfective and imperative verbs are expressed by the pronominal suffixes listed in §6.2.2. The addition of these suffixes in many forms results in resyllabification of the base of the verb. The paradigm of the perfective form of the verb ‘to kill’ (*gəṭal*) with the 3ms. (-i) and the 2pl. (-*kon*) object suffixes is given in Table 13 (after Häberl 2009: 181). In this paradigm there is phonological gemination (§5.6.2.) of the /n/ before the 3ms. object suffixes attached to 3pl. and 2pl. verbs.

Table 13: Pronominal objects with perfective and imperatives in Neo-Mandaic

3ms.	<i>gəṭl-i</i>	<i>gəṭal-kon</i>
3fs.	<i>gəṭal-t-i</i>	<i>gəṭl-at-kon</i>
3pl.	<i>gəṭl-onn-i</i>	<i>gəṭl-on-kon</i>
2s.	<i>gəṭal-t-i</i>	
2pl.	<i>gəṭal-tonn-i</i>	
1s.	<i>gəṭal-t-i</i>	<i>gəṭl-it-kon</i>
1pl.	<i>gəṭal-n-i</i>	<i>gəṭal-ni-kon</i>

The 1st and 2nd person imperfective forms of the verb express pronominal objects by object suffixes, but 3rd person imperfective forms express pronominal objects by oblique suffixes consisting of an originally dative preposition *l-* and pronominal suffixes, analogous to the L-suffixes of NENA. The paradigm of the imperfective realis form of the verb ‘to kill’ (*qə-gātel*) ‘he kills’ with the 3ms. (-*i*, -*l-i*) and the 2pl. (-*kon*, *lə-kon*) object suffixes is given in Table 14 (after Häberl 2009: 182).

Table 14: Pronominal objects with imperfective verbs in Neo-Mandaic

3ms.	<i>qə-gatel-l-i</i>	<i>qə-gatel-lə-kon</i>
3fs.	<i>qə-gatl-ā-l-i</i>	<i>qə-gatl-ā-l-kon</i>
3pl.	<i>qə-gatl-əl-l-i</i>	<i>qə-gatl-əl-kon</i>
2s.	<i>qə-gatl-āt-i</i>	—
2pl.	<i>qə-gatel-tonn-i</i>	—
1s.	<i>qə-gatel-n-i</i>	<i>qə-gatel-nā-kon</i>
1pl.	<i>qə-gatel-nann-i</i>	<i>qə-gatel-nan-kon</i>

In this paradigm there is phonological gemination (§5.6.2.) of the /n/ before the 3ms. object suffixes attached to 1pl. and 2pl. verbs. The /n/ of the 3pl. plural subject inflection -*en* assimilates totally the // of the oblique suffix.

Unlike most NENA dialects, Neo-Mandaic has not developed innovative verbal forms to express the present perfect, but rather uses the suffix conjugation form with this function in addition to the function of expressing the past perfective. The present perfect function includes resultative and stative (Häberl 2009: 240), e. g.

(12) Khorramshahr

mekt-at *əšta*
 die.PST-3FS. now
 ‘Now she is dead.’

The imperfective form expresses a variety of subtypes of imperfectivity, including progressive, habitual and narrative present. It also expresses the future (Häberl 2009: 241–243).

The Neo-Mandaic indicative copula has a short and a long form. The short form consists of clitics which are in origin light clitic forms of personal pronouns. Häberl (2009: 230) found the paradigm of this to be defective in the Khorramshahr dialect, with no short forms being in use for the 1st and 2nd plural forms, and no distinctions in gender made in any person:

3s.	= <i>ye</i>	2pl.	—
3pl.	= <i>non</i>	1s.	= <i>nā</i>
2s.	= <i>yāt</i>	1pl	—

Macuch (1993: 53) presents a slightly fuller paradigm for the Ahvāz dialect:

3ms.	=ye ~ =ya	2s.	=yāt ~ =yet
3fs.	=i	2pl.	—
3mpl.	=nɔn, =non	1s.	=nɔn
3fpl.	=nen	1pl.	—

The long copula, which bears its own stress, is formed by combining the pronominal suffixes listed in §6.2.2 with the base of the existential particle *eł-*. This results in phonological gemination of the interdental /t/, which, by the process discussed above in §5.6.2., has the outcome /kt/. The paradigm of the long copula is complete, with 1pl. and 2pl. forms and with gender distinctions in the 2nd and 3rd persons singular (Häberl 2009: 231):

3ms.	<i>ełt-i</i>	2fs.	<i>ełt-ek</i>
3fs.	<i>ełt-a</i>	2pl.	<i>ełt-əkon</i>
3pl.	<i>ełt-u</i>	1s.	<i>ełt-e</i>
2ms.	<i>ełt-ak</i>	1pl.	<i>ełt-an</i>

In addition to the preverbal particle *qə-*, which is used productively to express the realis of the imperfective, the non-productive preverbal particle *d-* can be identified in the imperative form of the verb *ałā* ‘to come’, viz. *doti!* ‘come!’ (Häberl 2009: 191; Macuch 1965: 306). This appears to be the same particle that is used before imperatives in various NENA dialects to express intensity.

8. Syntax of basic clauses

8.1. Copula clauses

8.1.1. NENA dialects

The syntax of copula clauses in NENA has been described in chapter 3.4, §8.1. Here we shall only make the observation that both in the Jewish dialects of the region and in C. Sanandaj the copula is fixed on the predicate and cannot be moved onto other constituents of the clause when they are in focus. This feature of the syntax of the copula is found also in the trans-Zab Jewish dialects of Iraq and northwestern Iran. The lack of movement is illustrated in the following examples:⁶

⁶ Intonation group boundaries are expressed by the symbol ‖ and nuclear stress by a grave accent.

- (13) J. Sanandaj (Khan 2009: 336)

tat-óx tajér=ye? *mam-i tajér=ye* *lá*
 father-your merchant=COP.3MS uncle-my merchant=COP.3MS NEG
tat-i.
 father-my

‘Is your father a merchant?’ ‘*My uncle* is a merchant, not my father.’

- (14) C. Sanandaj

báb-ox tajér=ile? *'axón-i tajér=ile* *lá*
 father-your merchant=COP.3MS brother-my merchant=COP.3MS NEG
báb-i.
 father-my

‘Is your father a merchant?’ ‘*My uncle* is a merchant, not my father.’

The copula is negated by combining it with the negative particle *la*, e. g. J. Sanandaj =*y(e)* ‘he is’, *la=y* ‘he is not’. In C. Sanandaj the // of the 3rd person copula forms is elided in the negative, e. g. =*ile* ‘he is’, *la=y* ‘he is not’.

8.1.2. Neo-Mandaic

The enclitic short copula of Neo-Mandaic is most commonly used in clauses in which the predicate is an adjective (15), a resultative participle (16), a locative expression (17) or interrogative expression (18) (Häberl 2009: 230, 253):

- (15) Khorramshahr

šəbir=nā
 good=COP.1S
 ‘I am well’

- (16) Khorramshahr

yehəm=ye
 sit.PTCP=COP.3S
 ‘He is seated’

- (17) Khorramshahr

ahni barra əm-belāt=non
 they outside from-city=COP.3PL
 ‘They are outside the city’

- (18) Khorramshahr

man=ye?
 who=COP.3S
 ‘Who is he?’

A clause with a short enclitic copula is negated by the part *lu* placed before the predicate:

- (19) Khorramshahr (Häberl 2009: 291, IX.10)

bāb-e lu-hnā=ye
father-my NEG-here=COP.3S
'My father is not here'

The long copula is regularly placed after the predicate. It may be used with the types of predicate exemplified in (15)–(18) and also with predicates containing a nominal, where it is the norm (Häberl 2009: 231), e. g.

- (20) Khorramshahr

an qanāyā d-kaspā ekt-e
1sg smith of-silver COP-1SG
'I am a silver smith'

The long copula is negated by prefixing the negative particle *la* to the base of the copula, the vowel of which is elide, e. g.

- (21) Khorramshahr

šəbir l-ekt-e
good NEG-COP-1S
'I am not well'

Copula clauses that are not indicative present are constructed with a form of the copula verb *həwā*, i. e. *həwā* past, *hāwi* irrealis, *qə-hawi* imperfective progressive/future.

8.2. Verbal clauses

8.2.1. NENA dialects

In the Jewish dialects of the region the direct object constituent of a clause is normally placed before the verb, as in the Jewish trans-Zab dialects of Iraq and northwestern Iran. This is also the normal word order in C. Sanandaj, which is noteworthy since in other regions the Christian dialects are generally distinguished from the Jewish trans-Zab dialects in having a basic verb–object word order.

In all dialects of the region a noun phrase that expresses a goal is normally placed after the verb that expresses movement, e. g.

- (22) J. Sanandaj (Khan 2009: 358)

zil-ex bela
go.PST-D.1PL house
'We went home.'

- (23) C. Sanandaj (Panoussi 1990: 120)

'az-ən besa
go.PRS-D.1S house
'Let me go home.'

All verbs forms, including imperatives, are negated by the particle *la*, e. g.

- (24) J. Sanandaj (Khan 2009: 92)

la gärəš-Ø
NEG pull.PRS-D.3MS
'He does not pull.'

la gruš!
NEG pull.IMPER
'Do not pull!'

In the Jewish dialects when *la* negates a perfect formed by the combination of a resultative participle and a copula auxiliary, the negator is placed before the participle rather than the auxiliary, e. g.

- (25) J. Sanandaj (Khan 2009: 97–98)

la smiqa=y
NEG redden.PTCP=COP.3MS
'It has not become red.'

If, however, the predicate is an adjective, the negator is placed before the copula:

- (26) J. Sanandaj (Khan 2009: 97–98)

smoqa la=y
red NEG=COP.3MS
'It is not red.'

8.2.2. Neo-Mandaic

In Neo-Mandaic the position of the direct object constituent depends to a large extent on its specificity (Häberl 2009: 135–137). Objects that refer to generic classes or indefinite objects, marked with the indefinite suffix *-i*, with non-specific referents are generally placed before the verb:

- (27) Khorramshahr (Häberl 2009: 135)

šam[‘] qə-masrik-en
candle IND-kindl.IPFV-3PL
'They light candles.'

- (28) Khorramshahr (Häberl 2009: 135)

jisr-i tum əbad-yon
 bridge-INDF again do.PST-3PL
 ‘They built another bridge.’

Compound verb constructions with light verbs (§7.2.), which replicate the pattern of Persian models, regularly place the generic nominal element before the verb, e. g.

- (29) Khorramshahr

hitra əbad
 celebration make.PST.3MS
 ‘he celebrated’

A definite object with an identifiable referent is generally placed after the verb and is introduced by the preposition *əl-* (‘to/for’). There is also object agreement on the verb in the form of an anticipatory object suffix, though this is elided before the object marker and its original presence is often only discernible in the syllabic structure of the verb (e. g. the phonological gemination of the /n/ of the 3pl. subject suffix in 30):

- (30) Khorramshahr (Häberl 2009: 136)

əbd-onn-Ø əl-ā konførens
 do.PST-3PL-OBJ.3MS OBJ-this conference
 ‘They held this conference.’

The specific direct object of a compound verb is generally marked by the preposition *qam* ‘to/for’. In such constructions there is no object agreement on the verb:

- (31) Khorramshahr (Häberl 2009: 137)

čāre abd-etton ... qam tā jisər
 remedy do.IPFV-2PL to that bridge
 ‘You will fix that bridge.’

The constituent expressing the goal of an action is regularly placed after the verb, e. g.

- (32) Khorramshahr (Häberl 2009: 278, III, 6)

ak waxt ezg-ini qam Haft Tappé
 that time go.PST-1PL to Haft Tappe
 ‘At that time we went to Haft Tappe.’

Perfective and imperfective forms of verbs are negated by the particle *la*, which bears the main word-stress:

- (33) Khorramshahr (Häberl 2009: 81)

lá-qə-mhatten
 ‘they do not talk’

Unlike the NENA dialects of the region, imperative forms cannot be negated in Neo-Mandaic. Negative commands and prohibitions are expressed by combining the negative particle *la* with the irrealis imperfective, e. g.

- (34) Khorramshahr (Häberl 2009: 247)

l-all-etton ger welāt
NEG-go.IRR-2PL outside city
'Do not go outside the city.'

9. Syntax of clause linkage

9.1. Coordinating linkage

9.1.1. NENA dialects

The means of coordinating linkage of main clauses in the NENA dialects of the region corresponds to the typology of the dialects of Iraq (chapter 3.4, §9.1). It may be asyndetic or marked by a particle. Asyndetic linkage is particularly common in the dialects of this region and appears to be the norm in C. Sanandaj judging by the texts published by Panoussi (1990).

In the Jewish dialect main clauses are occasionally linked by the co-ordinating particle *u*. This is generally attached as an enclitic to the last item of a clause before an intonation group boundary.

- (35) J. Sanandaj (Khan 2009: 368)

xa-yoma zil-Ø lag-ef=u mir-e baq-ef
one-day go.PST-D.3MS to-PRO.3MS=and say.PST-L.3MS to-PRO.3MS
'One day he went to him and said to him.'

Another strategy of coordinating linkage is to place a particle marking topical prominence on the subject of the linked clause:

- (36) J. Sanandaj (Khan 2009: 372)

'uč-lu 'e-jəl-ăke 'ana-č 'asr-ăna-nu
trample.IMP-L.3PL DEM-clothes-DEF I-TOP wring.PRS-D.1FS-L.3PL
'Trample the clothes and I shall wring them out.'

9.1.2. Neo-Mandaic

Coordinating linkage in Neo-Mandaic may be marked by the coordinating particle *u* (37) or it may be asyndetic (38):

- (37) Khorramshahr (Häberl 2009: 280, IV.19)

inšānā aṭon u qāren əlāw mienā
women come.PST.3PL and recite.IPFV.3PL to water
'Women came and enchanted the water.'

- (38) Khorramshahr (Häberl 2009: 290, VIII.6)

Hafté horettā aṭā daš
week other come.PST.3MS enter.PST.3MS
'He came another week and entered.'

9.2. Subordinating linkage

9.2.1. NENA dialects

The general subordinating particle *d* of earlier Aramaic has been lost before subordinate clauses in the dialects of the region.

Relative clauses are normally asyndetic in C. Sanandaj, without explicit linkage between the clauses and their heads:

- (39) C. Sanandaj

darmane doktor tm-ewe-lu
medicines doctor PST-give.PRS.D.3MS-L.3PL
the medicines that the doctor has given

Occasionally the Iranian particle *ke* is used to introduce the relative clause:

- (40) C. Sanandaj

'ay-xamra ke t-mand-ox-le
this-wine which PST-throw.PRS-D.1PL.-L.3MS
this wine which we have pressed

In J. Sanandaj asyndetic relative clauses are mainly restricted to cases where the head noun is indefinite:

- (41) J. Sanandaj (Khan 2009: 381)

xa-t^wka qārira hāw-e
a-place cool be.PRS-3MS
'a place that is cool'

When the head noun is definite the relative clause is generally introduced by the Iranian particles *ya* or *ke*:

- (42) J. Sanandaj (Khan 2009: 377–378)

'o-naše ya-da 'wat kol-i-wa-lu
the-people REL-invitation make.PRS-D.3PL-PST-L.3PL
'the people whom they invited'

- (43) J. Sanandaj (Khan 2009: 377–378)

xa-məndix=ye ke pərče koma ko-lu.
a-thing=COP.3MS REL hair.PL black make.PRS-L.3PL
'It is a thing that makes hair black.'

Factive and irrealis content clauses that are direct complements of verbs are sometimes introduced by the Iranian particle *ke*:

- (44) J. Sanandaj (Khan 2009: 386–388)

hāmər-Ø ke 'ay-brona həl-day brata gb-e.
say.IRR-D.3MS COMP that-boy OBJ-DEM.OBL girl love.PRS-D.3MS
'[in order that] he say that the boy loves the girl'

- (45) J. Sanandaj

'ijaza hulmu ke- 'axni xlula hol-ex
permission give.IMPER.PL COMP-we wedding make.IRR-D.2PL
'Give permission for us to hold the wedding.'

- (46) C. Sanandaj

sʃareš kewəl-Ø ke 'ana gi-se-li
message give.PRS-D.3MS COMP I PERF-come.PST-L.1S
'He gives a message that "I have come."'

Nouns functioning as adverbials at the head of a subordinate clause are often linked to the clause by the Iranian *ezafe* particle:

- (47) J. Sanandaj (Khan 2009: 394)

waxt-e xlula wi-li
time-EZAFE wedding do.PST-L.1S
'when I married'

- (48) C. Sanandaj

waxt-e hak-i
time-EZAFE speak.PRS-D.3PL
'when they speak'

The Aramaic conditional particle has been completely replaced by the Iranian particle *'ağar*:

- (49) J. Sanandaj (Khan 2009: 396)
- 'āgar xast=et 'at gbe gānet.
if tired=COP.2MS you must sleep.PRS-D.2MS
'If you are tired, you must sleep.'
- (50) C. Sanandaj
- 'āgar 'awa 'as-e 'ana k-az-ən
if he come.PRS.IRR-D.3MS I IND-go-D.1MS
'If he comes, I shall go.'

9.2.2. Neo-Mandaic

In Neo-Mandaic the Aramaic subordinating particle *d* is still used in noun phrases to introduce genitive attributes, but it is no longer used to introduce subordinate clauses.

Relative clauses use relative particles that have been borrowed from contact languages, viz. *ke* ‘that’ from Persian and *illi* ‘which’ from Arabic. The particle *ke* is used to introduce restrictive relative clauses, e. g.

- (51) Khorramshahr (Häberl 2009: 166)
- att-on barnaš-ānā ke šibih-ānā kābš-el-l-u
bring.IMP-PL person-PL REL demon-PL subdue.IPFV-3PL-OBJ-3PL
'Bring the people who will subdue the demons.'

In some cases the head noun has the Persian ‘specifying’ -*i* suffix, which is borrowed from Persian relative clause constructions. This is seen in (52), which also exemplifies the use of an object resumptive pronoun to express the object function of the referent of the head noun in the relative clause:

- (52) Khorramshahr (Häberl 2009: 263)
- ezg-it dukk-ān-i ke həz-it-u awwáл
go.PST-1SG place-PL-i REL see.PST-1SG-3PL before
'I went to the places which I saw before.'

In the Khorramshahr dialect the particle *illi* introduces non-restrictive relative clauses with a definite head noun, e. g.

- (53) Khorromshahr (Häberl 2009: 165)
- q-abi-n amār-Ø genz farwāh Profesər Buckley
IND-want.IPFV-1SG say.IRR-1SG many thanks Professor Buckley
- illi genz əbād abd-at qam Manday-ānā
REL much work do.PST-3FS for mandaean-PL

'I want to give many thanks to Professor Buckley, who has done so much work for the Mandaean.'

In the Ahvāz dialect the Persian relative particle *ke* is used in such non-restrictive relative constructions (Häberl 2009: 260).

When the head noun of a non-restrictive relative clause is indefinite with a specific referent, there is no relative particle, e. g.

- (54) Khorramshahr (Häberl 2009: 164)

gāw Aḥwāz həwā tarmid-i ešm-i šiex ‘Abdalla
in Ahvāz be.PST.3MS tarmida name-3MS sheikh Abdallah
'In Ahvāz there was a tarmida,⁷ whose name was Sheikh Abdallah.'

Adverbial clauses are introduced by a preposition or are asyndetic. The verb is in the irrealis or realis form according to whether the event has been realized or not, e. g.

- (55) Temporal clause, Khorramshahr (Häberl 2009: 248)

qamāy malp-āt-i ana gatel-nā-kon geš
before teach.IRR-2SG-3MS 1SG kill.IRR-1SG-OBJ.2PL all
'Before you teach him, I will kill you all!'

- (56) Causal clause, Khorramshahr (Häberl 2009: 249)

at-on gāw Irān qə-beyy-en komak
come.PST-3PL in Iran IND-want.IPFV-3PL help
abd-əl-l-Ø əl-Rusya
do.IRR-3PL-OBJ-3FS OBJ-Russia
'They came to Iran (because) they wanted to help Russia.'

In conditional constructions the Persian conditional particle *agar* is used (Macuch 1965: 432; Häberl 2009: 250), e. g.

- (57) *agar pərāhā eh-l-e turt-i qə-zaben-nā*
if money COP-to-1SG COW-INDF IND-buy.IPFV-1SG
'If I have money, I shall buy a cow.'

10. Concluding Remarks

The NENA dialects of western Iran represent the far eastern periphery of the NENA dialect area. They exhibit a number of features that are not characteristic of the vast majority of NENA dialects in other regions. This applies in particular to the cluster of Jewish dialects. One of the most conspicuous differences between the Jewish dialects and the rest of NENA is their split-S alignment in clauses with verbs formed from past bases. In Khan (2017) I have argued that this development is likely to be an innovation within NENA, triggered by a greater degree of

⁷ The first grade of the Mandaean priesthood.

convergence to the ergative alignment of Kurdish. The use of D-suffixes with the past base is likely to have begun as a means of expressing the stative-type resultative perfects in maximally stative unaccusative verbs, as is found in some other peripheral NENA dialects, such as J. Urmi and C. Hertevin (chapter 2.5, §7). In the Jewish dialects of western Iran this would have developed further and spread to the expression of the perfective. The development of intransitive verbal forms in Turoyo and Mlahso with the inflection of D-suffixes on past bases should be considered to be an independent development.

The lexicon of the Jewish dialects and C. Sanandaj contains many loanwords from Kurdish and Persian. The available material on C. Sanandaj dialect also exhibits a number of loanwords from Arabic. The core vocabulary remains predominantly Aramaic in origin, but even this exhibits some loanwords (Khan 2009: 401–409).

Neo-Mandaic, which is a separate subgroup of Neo-Aramaic, exhibits numerous differences from the NENA dialects of Western Iran in phonology, morphology, syntax and lexicon. In a variety of features Neo-Mandaic is more archaic than NENA.

The morphology of the pronouns in Neo-Mandaic, for example, is more conservative in some respects than the NENA dialects of the region, such as its preservation of distinct forms for anaphoric third person personal pronouns (§6.2.2.).

The morphology of the verb in Neo-Mandaic (§7.2.) is more archaic in its profile than the verbal systems of NENA dialects, preserving, for example, the original perfective suffix conjugation and a wider range of derivative verbal forms and lacking innovations such as distinct forms for the present perfect. Verbal clauses in Neo-Mandaic, moreover, exhibit uniformly accusative alignment.

The form of the Neo-Mandaic copula appears to be more archaic than NENA. Parts of the paradigm of the short copula still have the resemblance of a clitic pronoun, the historical origin of the copula. The long copula, formed by the combination of the existential particle and pronominal suffixes, is a common feature of earlier Aramaic dialects, but does not exist in NENA.

The expression of prohibitions by negation of the irrealis imperfective rather than the imperative (§8.2.2.) is also more in line with earlier Aramaic dialects than with the NENA dialects of the region, which negate the imperative.

The core vocabulary of Neo-Mandaic is predominantly of Aramaic origin. Neo-Mandaic and NENA share a number of lexical items that are not found in other Neo-Aramaic subgroups (Mutzafi 2014: 117–143). Although the Neo-Mandaic communities in south-western Iran are now not in contact with NENA speakers further north, some lexical isoglosses between the two dialect groups may have spread at an earlier period by areal diffusion. This applies in particular to the common NENA word *məndi*, which appears to be of Mandaic origin (§6.3.2.).

11. Glossed texts

11.1. North-Eastern Neo-Aramaic

J. Sanandaj dialect

Winter (Khan 2009: 430)

An acute accent (́) marks non-nuclear stress. A grave accent (̀) marks the nuclear stress of an intonation group. A vertical line (|) marks an intonation group boundary.

- (58) *'āwālán t^wkáne là 'asr-í-wa.¹* *našé*
 first shops NEG close.PRS-D.3PL-PST people
g-ez-í-wa wárya ba-talgà-č.¹ *xor-t^wkané là*
 IND-GO.PRS-D.3PL-PST outside in-snow-also indeed-shops NEG
'asr-í-wa-lu.¹ *wa-ba- 'álawà-č¹* *rás̄m*
 close.PRS-D.3PL-PST-L.3PL and-in-addition-also custom
ga-dokà=ye-le¹ *másálán ga- 'axér-e*
 in-there=COP.PST-L.3MS for-example at-end-EZ
páyiz qárwa sətwà¹ hár xanáwadé ta-nòš-ef.¹ *xa-dána*
 autumn close winter each family for-self-POSS.3MS
hár məšpahá ta-nòš-ef.¹ *g-ezàl-Ø-wa¹* one-single
tórtá šákàl-Ø-wa-la.¹ *k-m-e-wa-l-ó*
 cow buy.PRS.-D.3MS-PST-L.3FS IND-bring-D.3MS-PST-L.3FS-PVP
ga-béla nòš-ef.¹ in-house self-POSS.3MS

First of all, they did not close shops. People would go outside even in the snow. Indeed they did not close the shops. In addition, there was a custom there, for example at the end of autumn just before winter⁸ each family, each family went in their turn and bought a cow. They would bring it back to their home.

- (59) *ráb-an k-é-wa dábəh-Ø-wa-l-ó*
 rabbi-POSS.1PL IND-come.PRS.3MS-PST slaughter-D.3MS-PST-L.3FS-PVP
baq-èf.¹ támíz k-òl-Ø-wa-la.¹ *xáé-wa*
 for-POSS.3MS clean IND-do.PRS-D.3MS-PST-L.3FS see.PRS.3MS-PST
tărefá lá-hawy-a,¹ *xáràb lá-hawy-a.¹* *'o-waxtára*
 unclean NEG-be.IRR-D.3FS polluted NEG-be.IRR-D.3FS that-time
qăšàng¹ 'o-năqolé-u *măqolè-u¹* *'onyexáe yá*
 beautiful DEM-unclean-and unclean-and those REL

⁸ Literally: close to winter.

gdidim⁹=yen| xārāb=yen,| lá k-əxl-i-lù|
sinews=COP.3PL polluted=COP.3PL NEG IND-eat.PRS-D.3PL-L.3PL
doq-Ø-wa-lú-u šolat-Ø-wa-lù.|
seize.PRS-D.3MS-PST-L.3PL-and throw-D.3MS-PST-L.3PL

Our rabbi would come and slaughter it for them. He cleaned it. He checked lest it be ritually unclean, lest it be polluted. Then he carefully took out what had to be removed, those parts that were sinews and were polluted, which they did not eat, and threw them away.

- (60) *pəsr-áké k-wál-Ø-wa-le ba-'ilá máre bel-ákè.|*
meat-DEF IND-give.PRS-D.3MS-PST-L.3MS in-hand master house-DEF
máre bel-áké mən-dáy pəsrà| mätú-Ø-wa-le
master house-DEF from-that.OBL meat put.PRS.3MS-Ø-PST-L.3MS
ga-tänurá bāšəl-Ø-wa-le.| k-əmr-i-wa-le
in-oven cook.PRS-D.3MS-PST-L.3MS IND-say-D.3PL-PST-L.3MS
qawurmá.| nät-èn-wa-le| ga-tʷká qärirá
preserved_meat take.PRS-D.3PL-PST-L.3MS in-place cool
mät-i-wa-le.| yaxčál lit-wa-lan
place-D.3PL-PST-L.3MS fridge there_is_not-PST-L.1PL
xor-’o-waxtārá.| mät-i-wa-le ga-xá-tʷka qärirà
still-that-time place-D.3PL-PST-L.3MS in-one-place cool
hăwé-Ø| péš-Ø ta-sətwà.|
be.IRR-D.3MS stay.IRR-D.3MS for-winter

He then gave the meat to the head of the household. The head of the household put out some of this meat and cooked it in an oven. They called this *qawurma*. They took it and placed it in a cool place. We still did not have a fridge at that time. They put it in a place that was cool for it to remain until winter.

- (61) *’áy qawurmá ’äxà=ye-le| galgl-i-wa-le*
this preserved_meat thus=be.PST-L.3MS stir.PRS-D.3PL-PST-L.3MS
ga-pliyáw məšxá nòš-ef,| məšxá tort-ákè,| lá məšxà,
in-middle oil self-POSS.3MS oil cow-DEF NEG oil
denulà.| ga-denulá nòš-af bašl-i-wa-la
fat in-fat self-POSS.3FS cook.PRS-D.3PL-PST-L.3FS
qäšäng.| mät-i-wa-le ga-xá ... ó waxtāra
well put.PRS-D.3PL-PST-L.3MS in-one that time
qoq-è- hit-wa.| qoq-é xa-zárf,
pot-PL there_is-PST pot-PL one-container

⁹ Hebrew.

zrúf-ek=ye-lu *mán* *'ápr-e* *smoqà.* | *'ənyexáe*
 containers-INDF=be.PST-L.3PL from clay-EZ red they
mát-i-wa-lu *ga-sətwà.* | *mən-dóá* *rábá*
 place.PRS-D.3PL-PST-L.3PL in-winter. from-that.OBL much
'əstəfáda *k-ol-i-wa.* |
 advantage IND-do.PRS-D.3PL-PST

This *qawurma* was like this, they would stir it round in its own oil, the oil of the cow. Not oil, fat, they cooked it well in its own fat. They put it in a—at that time there were pots. Pots were a container, containers made of red clay. They would serve this in winter. They made good use of this.'

11.2. Neo-Mandaic

Khorramshahr dialect

The cooking pot (Häberl 2009: 292)

- (62) *ya yum-i ezgā-Ø Joha təlab-Ø*
 one day-INDF go.PST-3MS Joha ask.PST-3MS
əm-šəbāb-ān-i qazgān-i tāpi-Ø
 from-neighbour-PL-POSS.3MS cooking_pot-INDF cook.IRR-3MS
gāww-i
 in-POSS.3MS

'One day, Joha went to ask his neighbours for a cooking pot in order to cook in it.'

- (63) *ba 'ad so ruz kammar-Ø, bədaq-Ø kars-i*
 after three day return.PST-3MS put.PST-3MS belly-POSS.3MS
qazgān honin-i
 cooking_pot small-INDF
 After three days he returned, having put a small pot inside it.

- (64) *šəbāb-ān-i həz-onn-(i)¹⁰ əl-qazgān u*
 neighbour-PL-POSS.3MS see.PST-3PL-3MS OBJ-cooking_pot and
əkad-Ø-i minjo-d-i
 take.PST-3MS-3MS from_within-GEN-3MS
 His neighbours saw the pot, and he took it out of it.

¹⁰ Brackets enclose suffixes that have been elided in sandhi.

- (65) *mal-lonn-(i)* *əl-Joha* *mu=yye* *āhā?* *Joha* *əmall-u*
say.PST-3PL-3MS to-Joha what=COP.3MS this Joha say.PST-3PL
āhā *qazgān-kon*
this cooking_pot-2PL
He said to Joha “What is this?” Joha said them “This is your cooking pot.”
- (66) *həwā-Ø* *qār-e* *tam-Ø* *qār-i* *jihel.*
be.PST-3MS by-1MS become.PST-3MS by-3MS child
ati-t-i *qam-də-kon*
bring.PST-1S-3MS for-GEN-2PL
While it was at my place, it gave birth to a child. I brought it for you.
- (67) *ṣāheb* *bietā* *əmall-Ø-i* *muḥāl=ye* *dəri-Ø*
owner house say.PST-3MS-3MS absurb=COP.3MS take.PST-3MS
u *ahbi-Ø*
and keep-3MS
The owner of the house said to him “That is absurd”. He took it and kept it.
- (68) *hafté* *horetta* *aṭā-Ø* *Joha* *u* *ṭəlab-Ø* *minn-i*
week other come.PST-3MS Joha and ask.PST-3MS from-3MS
ya *qazgān-i*
one cooking_pot-INDF
The next week Joha came and asked him for a cooking pot.
- (69) *hab-lonn-i* *ya* *tafšā* *rabb-i.* *dər(i)-Ø-(i)*
give.PST-3PL-3MS one stew_pot big-INDF take.PST-3MS-3MS
əl-rabbā. *lá-kamr-i* *qam-didu*
OBJ-big NEG-return-3MS to-GEN.3PL
They gave him a large stew pot. He took the large one. He did not return it to them.
- (70) *ba'ad* *esbu'a* *ezgā-Ø* *šəbābā* *qam* *Joha.*
after week go.PST-3MS neighbour to Joha
əmall-Ø-i *qazgān-d-e* *elli=ye?*
say.PST-3MS-3MS cooking_pot-GEN-1S where=COP.3MS
A week later, the neighbour went to Joha. He said to him “Where is my cooking pot?”

- (71) *əmall-Ø-i qazgān met-Ø. əmall-Ø-(i) əl-Joha*
 say.PST-3MS-3MS cooking_pot die.PST-3MS say.PST-3MS-3MS to-Joha
man qabul q-ābed-Ø? qazgān mujur
 who acceptance IND-do.PRS-3MS cooking_pot how
qə-māyet-Ø?
 IND-die.PRS-3MS
 He said “The cooking pot has died.” He said to Joha “Who believes that a
 cooking pot could die?”
- (72) *Joha gəhek-Ø əmall-Ø-i man qabul q-ābed-Ø*
 Joha laugh-3MS say.PST-3MS-3MS who acceptance IND-do.PRS-3MS
qazgān tam-Ø qār-i štāna ‘ād qabul
 cooking_pot become.PST-Ø by-3MS boy then acceptance
lá-q-ābed-Ø mujur met-Ø?
 NEG-IND-do.PRS-3MS how die.PST-3MS
 Joha laughed, [and] said to him “Who would believe that a cauldron has
 given birth to a boy, but not believe that it has died?”

Abbreviations

Abbreviations not found in the Leipzig glossing conventions include:

D = D-suffix (i. e. direct, nominative pronominal suffix)

L = L-suffix (i. e. oblique pronominal suffix)

PVP = post-verbal particle

EZ = ezafe

References

- Borghero, Roberta. 2004. Some linguistic features of a Mandaean manuscript from the seventeenth century'. *Aram* 16. 61–83.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge Studies in Linguistics 70. Cambridge: Cambridge University Press.
- Häberl, Charles. 2009. *The Neo-Mandaic Dialect of Khorramshahr*. Semitica Viva 45. Wiesbaden: Harrassowitz.
- Häberl, Charles. 2011. Neo-Mandaic. In Stefan Weninger, Geoffrey Khan, Michael Streck & Janet Watson (eds.), *The Semitic Languages: An International Handbook*, 725–737. Berlin & Boston: De Gruyter Mouton.
- Heinrichs, Wolfhart. 2002. Peculiarities of the verbal system of Senaya within the framework of North Eastern Neo-Aramaic (NENA). In Werner Arnold & Hartmut Bobzin (eds.), “Sprich doch mit deinen Knechten Aramäisch, wir verstehen es!” 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag, 238–268. Wiesbaden: Harrassowitz.

- Hoberman, Robert. 1988. The history of the Modern Aramaic pronouns and pronominal suffixes. *Journal of the American Oriental Society* 108. 557–575.
- Hopkins, Simon. 1999. The Neo-Aramaic Dialects of Iran. In Shaul Shaked & Amnon Netzer (eds.), *Irano-Judaica IV*, 311–327. Jerusalem: Magnes.
- Hopkins, Simon. 2002. Preterite and perfect in the Jewish Neo-Aramaic of Kerend. In Werner Arnold & Hartmut Bobzin “*Sprich doch mit deinen Knechten Aramäisch, wir verstehen es!*” 60 *Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, 281–298. Wiesbaden: Harrassowitz.
- Israeli, Yaffa. 1998. *Ha-Aramit ha-hadashah shebe-fi Yehude Sakiz (derom Kurdistan)*. Jerusalem: Hebrew University dissertation.
- Kahn, Margaret. 1976. Borrowing and variation in a phonological description of Kurdish. Michigan: University of Michigan dissertation.
- Kalin, Laura M. 2014. Aspect and argument licensing in Neo-Aramaic’. Los Angeles: University of California dissertation.
- Khan, Geoffrey. 2004. *The Jewish Neo-Aramaic dialect of Sulemaniyya and Halabja*. Studies in Semitic Languages and Linguistics 44. Leiden & Boston: Brill.
- Khan, Geoffrey. 2009. *The Jewish Neo-Aramaic dialect of Sanandaj*. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2012. The evidential function of the perfect in North-Eastern Neo-Aramaic dialects. In Rebecca Hasselbach & Na’ama Pat-El (eds.), *Language and Nature. Papers presented to John Huehnergard on the occasion of his 60th Birthday*, 219–228. Studies in Ancient Oriental Civilization 67. Chicago: University of Chicago.
- Khan, Geoffrey. 2017. Ergativity in Neo-Aramaic. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *Oxford Handbook of Ergativity*, 873–899. Oxford: Oxford University Press.
- Macuch, Rudolf. 1965. *Handbook of Classical and Modern Mandaic*. Berlin & Boston: De Gruyter Mouton.
- Macuch, Rudolf. 1989. *Neumandäische Chrestomathie: Mit grammatischer Skizze, kommentierter Übersetzung und Glossar*. Porta Linguarum Orientalium. Neue Serie 18. Wiesbaden: Harrassowitz.
- Macuch, Rudolf. 1993. *Neumandäische Texte im Dialekt von Ahwāz*. Semitica Viva 12. Wiesbaden: Harrassowitz.
- Macuch, Rudolf & Estiphan Panoussi. 1974. *Neusyrische Chrestomathie*. Porta Linguarum Orientalium. Neue Serie 13. Wiesbaden: Harrassowitz.
- Mutzafi, Hezy. 2014. *Comparative lexical studies in Neo-Mandaic*. Studies in Semitic Languages and Linguistics 73. Leiden & Boston: Brill.
- Panoussi, Estiphan. 1990. On the Senaya Dialect. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 107–129. Atlanta: Scholars Press.
- Panoussi, Estiphan. 1991. Ein vorläufiges Verbgleich zum aussterbenden Neuaramäischen Senaya-Dialekt. *Rivista Degli Studi Orientali* 65(3/4). 165–183.
- Tsereteli, Konstantin. 1990. The velar spirant ǵ in Modern East Aramaic dialects. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 35–41. Harvard Semitic Studies. Atlanta: Scholars Press.

4.5. Hawrāmī of western Iran

Parvin Mahmoudveysi and Denise Bailey

1. Introduction

1.1. Background

Hawrāmī is spoken in the Hawrāmān (Persian: Ourāmān) region located in the northern Zagros mountains in northwestern Iran and northeastern Iraq. Hawrāmī is a variety of the language that is identified by the historical name of Gorani (or Gūrānī), traditionally classified as Northwestern Iranian. Gorani, in turn, can be described as a continuum of closely related varieties, which nonetheless forms what is essentially a single language. In Iran, Hawrāmī shares this continuum with the endangered varieties of Kandūlāyī, Zardayāna, and Gawrajūyī, as well as with others no longer in active use, such as Rījābī, Gahwāraī, and Bēwanījī.¹ In Iraq, Hawrāmī shares this continuum with Bājałānī, which is historically spoken near Sar Pol-e Zohāb but also spoken northwest of Mosul, as well as with the varieties identified as Šabakī, Kākayī, Šēxānī, Zangana, and Rožbayānī (see Haig, this volume, p. 297, for a map).² In comparison to the other Gorani varieties, Hawrāmī is considered to be a “more archaic variant in many ways” (MacKenzie 2002: n.p.) and “probably [...] the best preserved of the group” (MacKenzie 1966: 4). MacKenzie’s characterization is supported by the relatively complex morphology of Hawrāmī, which includes extensive use of grammatical gender, regular oblique case marking, the Imperfect, and two forms of the *ezāfe* particle.³

Much of known Gorani history has been in close association with Kurdish. Gorani had a highly significant role as the literary and court language of the Kurdish Ardalān dynasty from the fourteenth to nineteenth centuries (MacKenzie

¹ See Mahmoudveysi (2016: 3).

² Actual names for Gorani and varieties can differ, depending on speaker and location (see Leezenberg 1993: 7). In this study, we use the name ‘Gorani’ to refer to the general northwestern language and ‘Hawrāmī’ to refer to the specific cluster of varieties spoken in the Hawrāmān area (for more discussion of the identification of Gorani, see Mahmoudveysi, Bailey, Paul, and Haig 2012: 2–4).

³ Gorani has been the focus of scholarly research since at least the early 1800s, with Hawrāmī figuring in many studies, including Benedictsen and Christensen (1921), Mann and Hadank (1930), Christensen (1936), and MacKenzie (1961, 1966, 1987b, 2002), Blau (1989), Naqšbandī (1996), Paul (2007), Holmberg and Odden (2008), Mahmudweyssi and Haig (2009), Zolfaqari (2010), and Bistoon, Gheityur, and Kazzazi (2013).

2002: n.p.). They share many linguistic features that serve as evidence of contact and convergence from centuries of proximity.

Hawrāmī is mainly found within the national boundaries of Iran. It is used in the Hawrāmān region, which is located in Kermanshāh and Kurdistan provinces and situated between the cities of Marīvān, Sanandaj, and Pāwa (Persian: Pāveh). There are more speakers of Hawrāmī than of the other Gorani varieties in Iran, such as those in Zarda, Gawrajū, and Kandūla.⁴ Hawrāmī is also spoken in Iraq, where the western part of Hawrāmān includes the settlements of Tawēla and Bīyāra. Hawrāmī speakers live in Xurmāl and in larger Kurdish cities as well, including Halabjā.⁵

The Hawrāmān region is constituted of four main parts: 1) Hawrāmān-i Luhōn; 2) Hawrāmān-i Taxt; 3) Šāmyān and Dizlī; and 4) Hawrāmān Řazāw-u Kamara.⁶ There are also two other parts: 5) Hawrāmān Gāwaro; and 6) Hawrāmān Žāwaro (these might be a part of Řazāw-u Kamara).⁷ The first part, Hawrāmān-i Luhōn, has about thirty-four settlements in Iran, nine of which are populated by speakers of Jāfī, a dialect of Central Kurdish. The settlements in Hawrāmān-i Luhōn include the historical center, Nawsūd, and also the city of Pāwa and villages such as Hāna Garmāla, Kaymna, Dizāwar, Nodša (Notša), Hajj, Narwē, Šošmē, Qałagā, Darmūr, Xānagā, Girāł, Šēxān, and Naysāna. Settlements in Iraq include Darē Mař, Sargat, Gułp, Baxa Kon, Dagā Šēxān, Sosakān, Bīyāra, and Tawēla. The second part, Hawrāmān-i Taxt, consists of the center, Šār-i Hawrāmān, and the villages of Waysīā, Kamāla, Sar-u Pīrī, Silēn, Biłbar, Žīwār, Kaljī, Nāwa, Nwēn, Dał, Dałamarz, Zom, and Asparēz, as well as Bānī Binok in Iraq. The third part, Dizlī (Dezlī) (and Šāmyān and Dizlī), includes the center, Dezlī, and the villages in Iran of Bārāmāwā, Zalka, Qałajē, Qałagā, Tiflīya, Tāzāwā, Goṛgayī, Tēžtēž, Tālawān, Xošāw, Xirosa, and Dara, and in Iraq, as Hāna, Qułē, Załm, and Ahmađ Āwā. The fourth part, Hawrāmān Řazāw-u Kamara, consists of the center, Řazāw, with the villages of Kařāwa, Dēwaznāw, Dagāgā, Xānagā, and Nāwa. The fifth and sixth parts, included by Hawrāmānī (2001), are Hawrāmān Gāwaro, with villages of Darwēšān, Galēn, Dē Molē, Hašamēz, Tawrēbar, Wasē, and Tāw Doław; and the area of Hawrāmān Žāwaro, with the villages of Māzībin, Borīdar, Čamšīdar,

⁴ For Gorani of Gawrajū, see Mahmoudveysi, Bailey, Paul, and Haig (2012); for Gorani of Zarda, see Mahmoudveysi and Bailey (2013); and for Gorani of Kandūla, see Mann and Hadank (1930).

⁵ For the locations of these varieties, see Haig (this volume, chapter 3.3, Fig. 4). Further information on locations is found in Mahmoudveysi (2016: 6–7) and MacKenzie (1966: 5). In Iraq, Gorani and its varieties are sometimes referred to collectively as ‘Māčo’ (or also ‘Hawrāmī’, as noted by Leezenberg 1993: 7).

⁶ See also MacKenzie (1987a), MacKenzie (1966: 5), Paul (2007: 287), and Zolfaqari (2010: 324) for parts.

⁷ See Hawrāmānī (2001: 617–619) for these areas and names of villages.

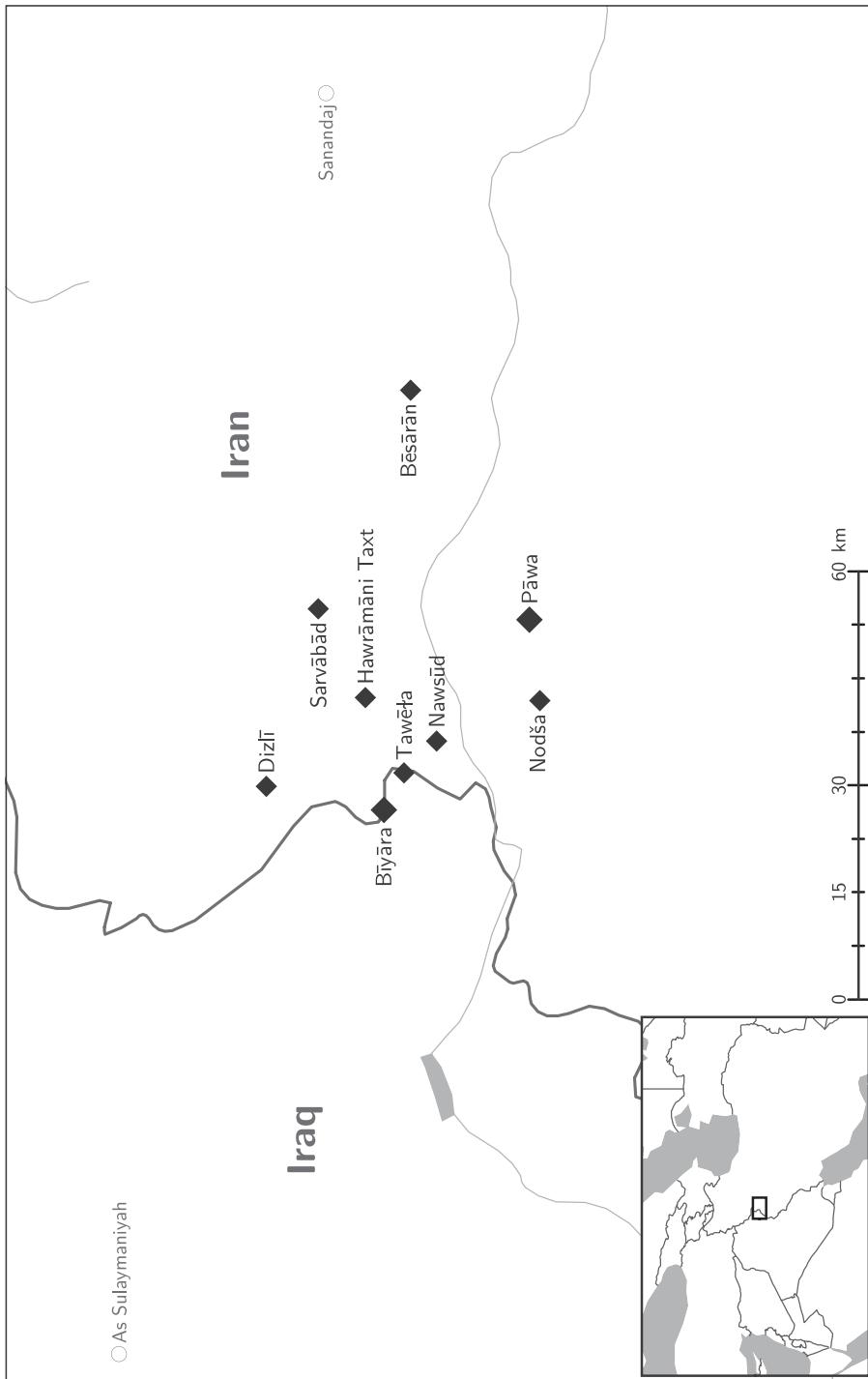


Figure 1: Locations of Hawrami mentioned in this chapter

Bēsārān, Žanīn, Tangībar, Pālingān, Žirēža, Hāna-w Hisayan Bagī, Sūra Tifē, Dēr Mołē, Kēlāna, and Takya.⁸

1.2. Names

The name *Hawrāmī* is used by people of the region to designate in general the speech of the settlements in Hawrāmān. More specific names include *Pāwayī*, used by people from Pāwa to refer to their own speech, in which case *Hawrāmī* is then used as a reference to the speech of Nawsūd.⁹

1.3. Literature

Gorani is used for most of the sacred literature of the Yārsān (Ahl-e Haqq) community. Gorani is especially known for its literature produced from the fourteenth to the nineteenth centuries, with its golden age from the seventeenth to nineteenth centuries. It was the literary language for all Kurdish and Gorani speakers in the entire area from Ilam to Kirkuk. The most prestigious of the Gorani poets, Sayyed Abdul Karim Tawgozī, also known as Mawlawī Kurd (1806–1882), was in fact primarily a Kurdish speaker who composed all of his poetry in Gorani.

Gorani literature developed and flourished during the Ardalan dynasty (1169–1867). After the downfall of the Ardalan rulers, Gorani declined in its position as the language of poetry, and Sorani Kurdish poetry gradually took on this role and gained prestige through the support of the rulers of Sulaymaniya. Gorani poets include Mullā Mustāfā Besārānī (1642–1701), Xānāy Qubādī (1700–1759), Saydī Hawrāmī (1784–1852), Ranjūrī (1750–1809), Ghulam Razā Arkawāzī (1775–1840), Mirzā Šafīq Jāmarēzī (1776–1836), Ahmad Bag Komāsī (1798–1878), Malā-y Jabbārī (1806–1876), and Sayyed Abdul Karīm Tawgozī, known as Mawlawī Kurd (1806–1882). There are many other Gorani poets; these names are only a selection of them.

⁸ A new administrative division of cities and villages in the Kurdistan province in Iran includes Sarvābād, consisting of Sarvābād and Hawrāmān Taxt. This also means that the Hawrāmān region is divided into two parts: 1) Hawrāmān Taxt (villages: Hawrāmān Taxt with Sar-u Pīrī, as a unit, Kamāla, Rēwar, and Wīsīan); and 2) Šālyār (villages: Silēn, Biłbar, Žīwār, Wargawīr, Nwēn, Kaljī, Nāw, ʕabāsābād, and Asparēz) (according to Sadeqi, n.d.).

⁹ Other local names for Hawrāmī spoken near Pāwa may be derived from the village name, for example, *Xānagāyī*, spoken in Xānagā, or *Dišayī*, spoken in Diša.

2. Background to the Hawrāmī speech community (Pāwa and Nawsūd)

The city of Pāwa (Persian: Pāveh) belongs to the county (*shahrestān*) of Pāveh in Kermānshāh province. It is located approximately 120 kilometers northwest of Kermānshāh. Pāwa is the capital; other towns include Nawsūd, Nodša, Bānawrē, and Bayangan. Hawrāmī is spoken in Pāwa and Nawsūd, while Jāfī Kurdish is spoken in Bānawrē and Bayangān. Jāfī is also spoken in most of the villages between Bānawrē and Pāwa, as well as in about four villages between Pāwa and Nawsūd. Hawrāmī is used in the remainder of the villages.

Recent estimates of Gorani speakers range from 200,000–300,000 (Paul 2007: 285) up to perhaps more than 500,000 (Mahmoudveysi 2016: 8). The ‘whole Avromani ethnic continuum’ is estimated at 80–90,000, with about 50,000 in Iran and about 30–40,000 in Iraq (Zolfaqari 2010: 238). The population of Pāwa and villages is estimated to be around 56,000; that of Nawsūd is estimated at 7,984; and that of Nodša as 3,547 (as of the 2006 census). The population of Sarvābād is about 70,000 people.¹⁰

Traditional daily life involved animal husbandry of sheep, goats, and cattle, with production of milk, yogurt, butter, and cheese. This work, however, has become less common in modern life. Most people tend gardens and orchards, growing walnuts, figs, apples, pomegranates, and grapes, as well as tomatoes and okra, which are ingredients in traditional Hawrāmī dishes such as *Doyna* and *Šalamīna*.

Most people in Pāwa are well educated, with an estimated literacy rate of 99 percent of the population between 10 and 49 years of age.¹¹ Men and women are trained in many types of professions and work in schools, clinics, and offices.

People in the Hawrāmān area primarily follow traditions of Sunni Islam. Hawrāmān is an important location for two Sufi (mystical Islamic) orders. Pāwa is one of the main locations for both orders, while the village of Nejār, close to Pāwa, is an important location for the Qāderīya order. The region also exhibits historical influence of other religions. One of the largest ancient fire temples of Zoroastrianism is found in Pāwa (Shahbazi, n.d.).¹² It is on a high mountain of the same name, Ātašgā.

2.1. Sociolinguistic situation

As mentioned above, the primary languages of this area are Hawrāmī and Jāfī, a variety of Central Kurdish. Hawrāmī is spoken in the cities and in most villages.

¹⁰ See Mahmoudveysi (2016: 7–8).

¹¹ See Paveh Press (2013), with this percentage estimated by Eqbal Zomorodi, Director of the Office for the Abatement of Illiteracy.

¹² In Gawrajū and Zarda villages, located south of Hawrāmān, Gorani speakers are of Yārsān (Yārī, Ahl-e Haqq) background. In Kandūla, the community is Twelver Shi‘ite (as noted by Paul 2007: 285).

It is used by an estimated 70 percent of the population of the Hawrāmān region in Pāwa, Nawsūd, and the villages between Pāwa, Nawsūd, and Marīwān. Jāfī is spoken in villages between Řawānsār and Pāwa, including Šimšēr, Darabāyān, Nosma, Čoržī, Bindara, and Dūrīsān, and also between Pāwa and Nawsūd, in the villages of Niğār, Noryāw, Wiřā, and Komara.

In terms of language use, both Hawrāmī and Jāfī are spoken within the family domain. Hawrāmī is normally spoken at home and learned by all children. Hawrāmī is used by speakers in work-related and marketplace domains as well. Hawrāmī has high prestige, and people have very positive attitudes towards it. In terms of multilingualism, it is observed that, generally, Jāfī speakers do not learn Hawrāmī, with rare exceptions. Most Hawrāmī speakers, however, are able to speak Jāfī. When Hawrāmī and Jāfī speakers converse, they generally choose to do so in Jāfī.

There are still a number of monolingual speakers of Hawrāmī, mostly women of the older generations. These monolingual speakers can nonetheless communicate to a degree with Jāfī speakers, with each participant using their own language.

There is no official education in Hawrāmī, though some young educated people have begun teaching Hawrāmī courses in Marīwān, Sanandaj, Pāwa, and the vicinity. A Hawrāmī festival takes place annually in Halabja or Sulaymaniya. Media include radio programs, and in recent years, an increasing number of local writers are producing literature in Hawrāmī. A magazine, ‘Hawrāmān’, is regularly published, and there are children’s books, short stories, and poetry, including a translation of Gulestan by the great Iranian poet Saadī.

Persian, as the official language of the Islamic Republic of Iran, is also spoken in the area. Modern Standard Persian is used in all official institutions, such as schools, hospitals, and the military, and also in the media. It is the language of wider communication for people who have come from other Iranian cities. It is learned by children in school and through exposure to television, but it is not typically used in the Hawrāmī family domain. Colloquial Persian is common in tourism and trade.

2.2. Present study

The rest of this study summarizes some of the linguistic features of Hawrāmī from Nawsūd and from Pāwa. The description of Hawrāmī of Nawsūd is based on the work of MacKenzie (1966). The description of Hawrāmī of Pāwa is based on material from two native speakers: Parvin Mahmoudveysi, one of the authors of this study, from Pāwa, forty-six years of age, who also speaks Kurdish, Persian, German, and English; and Jahangir Mahmoudveysi, also from Pāwa, fifty-six years of age, who also speaks Kurdish, Persian, and English. Jahangir Mahmoudveysi is a writer and author of poems and short stories, including the story presented at the end of this study, a contribution that is gratefully acknowledged.¹³

¹³ The authors would like to offer sincere appreciation to Jahangir Mahmoudveysi for his

In this description, language material can be identified as from Pāwa or from Nawsūd according to the sources. Pāwa material comes from the text by Jahangir Mahmoudveysi at the end of this study (bracketed numbers refer to these examples), from Mahmoudveysi and Bailey 2004, abbreviated as [M&B 2004], or from Mahmoudveysi 2017, abbreviated as [P.M. 2017]. Unspecified Pāwa material is also from Parvin Mahmoudveysi. Nawsūd material is from MacKenzie (1966), abbreviated as [D.M. 1966].

3. Phonology

The Hawrāmī material in this study is written with a mostly phonemic orthography.¹⁴ Material in this section on phonology is also transcribed in IPA (International Phonetic Alphabet; International Phonetics Association 2015/2005). The table below shows the correspondence of the orthographic symbols to the IPA symbols.¹⁵

Table 1: Orthographic symbols with IPA correspondences

Vowels		Consonants					
I	[i:]	p	[p ^h]	n	[n]	č	[tʃ ^h]
i	[ɪ-i]	b	[b]	l	[l]	j	[dʒ]
ē	[e:]	t	[t ^h]	ł	[ł]	x	[x]
e	[ɛ]	d	[d]	ř	[r]	gh	[ɣ] (loan)
a	[æ-ɛ-ɔ]	ð	[ð]	r	[r]	h	[h]
ā	[a:]	k	[k ^h]	f	[f]	h	[h]
ɔ	[ɔ]	g	[g]	s	[s]	y	[j]
ō	[o:]	q	[q]	z	[z]	w	[w]
u	[ʊ]	ł	[ł]	š	[ʃ]		
ū	[u:]	m	[m]	ż	[ʒ]		

contribution to this study and to Nicholas Bailey for his technical assistance and review of the grammatical description. Special thanks is also due to Laurentia Schreiber for her helpful editorial work.

¹⁴ Where relevant, Hawrāmī material from other sources (including MacKenzie 1966) is adapted to this orthography to facilitate comparison. The orthography also includes the symbol <d>, which, as an allophone, does not have full phonemic status, but it is still regarded here as important to signify.

¹⁵ Phonetic material is enclosed in square brackets, [...], phoneme symbols are enclosed by slash marks, /.../, and orthographic symbols are enclosed by brackets, <...>. Primary stress is marked with ['] before the syllable. Language material in the prose sections is in italic font.

3.1. Consonants

Table 2: Consonant phonemes (IPA)

	Bi-labial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p ^h b		t ^h d			k ^h g	q		
Affricate				tʃ ^h dʒ					
Fricative		f	s z	ʃ ʒ		x	(Pāwa: loan ㅂ)	h ՚	h
Nasal	m		n						
Trill			r						
Tap			t						
Lateral			l						
Lat. velarized			{						
Semivowel	w			y					

The voiceless plosives /p, t, k, q/ [p^h, t^h, k^h, q^h] have non-contrastive aspiration. A non-phonemic glottal plosive [?] is slightly audible before a word-initial vowel. The voiced plosives /b/ and /g/ maintain obstruent articulation in various environments.

The voiced plosive /d/ is of particular interest on account of the ways it is realized within the varieties in the Hawrāmān region and in the wider area. In Pāwa, /d/ maintains obstruent articulation in word-initial and post-consonantal positions, such as following /r/: *wirdī* ['wirdɪ] ‘small’ [M&B 2004: P85]. In post-vocalic environments, however, /d/ undergoes lenition. It is realized as what might be described as “a half-close central continuant caused by the tip of the tongue approaching the upper teeth without making contact” (MacKenzie 1966: 7–8). This sound in the Pāwa material seems to involve a slight quality of tongue-backing and with the lips slightly spread. It is symbolized here with ڏ [ð]. Examples include: *adā* [?ə'ða] ‘mother’ [M&B 2004: P31]; *badān* [bə'ðən] ‘body’ [M&B 2004: P135]; *didān* [dɪ'ðən] ‘tooth’ [M&B 2004: P19] (also [dɪ'ɪa]) [P.M. 2017]; and *ēd* [?eð] ‘PROX.3SG’ [M&B 2004: P3]. It seems to be close to that which is described as the alveolar approximant [ɹ] by Naqšbandī (1996: 125), who also notes devoicing at the end of a word, thus [ɹ], as in [mah'mʷuɹ] ‘Mahmoud’ (our stress transcription).

This phoneme maintains an obstruent articulation following /r/ not only in Pāwa, but also in the speech of the villages between Kāmyārān and Sanandaž. Moreover, it is found further south in the Gorani variety of Zardayāna, as illus-

trated by *bardan* [bær'dæn] ‘has taken’ (Mahmoudveysi and Bailey 2013: 150) and in the variety of Gawrajūyī, as illustrated by *wardē* [wær'de] ‘has eaten’ (Mahmoudveysi, Bailey, Paul, and Haig 2012: 94; with possible reduction or deletion word-finally or preceding a pause). In the village of Nawsūd, however, as well as in the villages between Pāwa and Nawsūd, extending to Marīvān, the sequence /rd/ is found as an approximant. For example, /rd/ in Pāwa *karday* ‘to do’ and *marday* ‘to die’ is found as /l/ in the corresponding words in the Nawsūd area.

The phoneme /d/ also undergoes lenition in the post-vocalic environment in Zardayāna, where it can be realized as [ð] or [j], for example, *mido* [mi'ðo] ‘gives’ and *ād* [?aj] ‘DIST.3SG’, or in vowel length, as in *ādam* [?a:m] ‘human being’ (Mahmoudveysi and Bailey 2013: 13). The realization as [j] is also found in the speech of Hawrāmī speakers outside of the Hawrāmān region, as illustrated by *ayā* [?æ'ja] ‘mother’, by speakers living in the city of Kermānsāh and in other cities outside of Hawrāmān, such as Karaj and Tehran.

The sound can also be found as a lateral approximant. Thus, the word pronounced by speakers in Pāwa as [χwo'ra] ‘God’ is pronounced by speakers outside of Hawrāmān in the city of Kermānsāh as [χwo'ra]/[χwo'la], as described by Naqšbandī (1996: 286, his transcription, our stress placement).

MacKenzie (1966: 8) mentions that /t/ in the second singular clitic pronoun =*it* can be realized as a ‘continuant’ in word-final position. This realization is also possible in Pāwa. The sound is provisionally symbolized aꝝ. (See (14), (20), (37) in our glossed text at the end of this study.)

The fricative series exhibits some complexities. The sound /x/ only has a voiced counterpart in loanwords, /v/. The phoneme /f/ also has irregular instances of a voiced counterpart. In Nawsūd and Nodšā, /v/ is an alternative for /w/ in word-initial position. In Pāwa, there is a sound that might be interpreted as [vʷ], as in *warwa* ['vʷərwa] ‘snow’ [M&B 2004: P68], or as this sound in free variation with [w]: *waratāw* [wərə'tʰaw] or [vʷərə'tʰaw] ‘sun at noon’ [M&B 2004: P55]; *wiñi* ['wiñi] or ['vʷiñi] ‘blood’ [M&B 2004: P21]; and *wirdī* [wir'di] or [vʷir'di] ‘small’ [M&B 2004: P85]. Some words beginning with /w/, however, are more consistently realized as such: *wārān* [wa'ran] ‘rain’ [M&B 2004: P67].

There are distinctions between a glottal fricative /h/ *hašt* [həʃtʰ] ‘eight’ [M&B 2004: P107] and a voiceless pharyngeal fricative: *haft* [həftʰ] ‘seven’ [M&B 2004: P106]; between a velarized lateral approximant /ɬ/ and simple /l/ (Pāwa): *gul* ‘flower’ and *gul* ‘dirty’; *kał* ‘mountain goat’ and *kal* ‘partially broken’; and between a trill /r/ and flap /t̪/: *haṛa* [hə'rə] ‘mud’ and *hara* ['hərə] ‘donkey’ [P.M. 2017].¹⁶ Neither /ɬ/ nor /r/ occur in word-initial position. The voiced pharyngeal fricative /ʕ/ is found mainly in loanwords from Arabic.

Preceding a pause (word-finally), /n/ and /g/ occur as a velar nasal [ŋ]: *māng*

¹⁶ The general term for ‘flower’ in Hawrāmī is *wiñi*, while *gul* is also found in literary Hawrāmī.

[maŋ] ‘moon’ [M&B 2004: P57]. A velar nasal can precede a velar plosive in word-medial position: *rangat* [rəŋgət] ‘your color’ (14).

There are only a few instances of gemination. In our Pāwa text, an example appears in *enna* ‘so much’ (22). Paul (2007: 289–290) also describes a particular type of nasal gemination developing from a combination of /n/ and /d/ in intervocalic environments.

3.2. Vowels

Table 3: Vowel phonemes (IPA)

	Front	Central	Back
Close	i i (P: i-i)	o	u
Close-mid	e ɛ		o
Open-mid	N: æ (æ-e), P: ə (ε-ə)		N: ɔ
Open		a	

As noted for Nawsūd by MacKenzie (1966: 8–9) and also found in Pāwa, the phoneme /i/ (MacKenzie’s transcription: *i*, our orthography: *i*) is often elided between consonants in syllables that do not bear the primary stress of the word: *wit* ‘he slept’ and *'nawt* ‘he did not sleep’ [D.M. 1966: 9, his gloss corrected here]. MacKenzie also notes that the phonemes /i/ and /u/ tend to be reduced in initial syllables in some contexts.

The vowel /ɔ/ in Nawsūd, but not Pāwa, has limited distribution, only occurring in the postposition =(*a*)*wa* after a consonant or vowels /i/ and /u/ (MacKenzie 1966: 10). This vowel /ɔ/ appears to correspond to the sequence /aw/.

In Nawsūd, /æ/ (MacKenzie’s transcription: *a*, also our orthography) is described as ranging from “an open front [a], normally” (that is, assumed here as IPA: [æ]) “to near half-open [ɛ]” (MacKenzie 1966: 9). In the Pāwa material, this vowel can also be described as having a range of phonetic realizations from [ɛ] to [ə].¹⁷

3.3. Stress

Stress is not entirely predictable, though it tends to occur on the final syllable of stems.¹⁸ It occurs regularly on morphemes marking plural, the demonstrative particles =‘*a* and =‘*ē*, subjunctive ‘*bi-*, and negation markers. It occurs on the final

¹⁷ Our analysis of vowel phonemes and the ranges of realization is provisional.

¹⁸ For discussion of stress in Hawrāmī, see Mahmoudveysi (2016: 76–80).

syllable of the definiteness markers. The position of word stress can mark a change in meaning, which can be illustrated by the shift to the first syllable of multisyllabic proper nouns to indicate vocative case (see also MacKenzie 1966: 21). For example: *Pa'rī* is realized as '*Parī*' in vocative case, similarly *Rē'bīn* as '*Rēbīn*' [P.M. 2017].

3.4. Common phonological processes

There is a tendency towards devoicing /t/, /z/, and /ʒ/ preceding a pause (word-final position): *ser* [ser / səɾ] ‘head’ [M&B 2004: P11]; *āyir* [?a'jɪɾ / ?a'jɪʃ] ‘fire’ [M&B 2004: P51]; *sawz* [səwz / səws] ‘green’ [M&B 2004: P78]; and *dirēž* [dɪ'reʒ / dɪ'reʃ] ‘long’ [M&B 2004: P91]. Voiced plosives tend to be devoiced and also unreleased preceding a pause: *zerd* [zəɾd / zəɾt] ‘yellow’ [M&B 2004: P79]; *serd* [səɾd / səɾt] ‘cold’ [M&B 2004: P88].

Devoicing is important in the formation of certain contracted forms whereby the initial consonant in the simple prepositions *ba* ‘to, with’ and *ja* ‘from, to, in’ is devoiced when combined with a postposition or other element. For example, the absolute preposition *pana* ‘to, by’ is actually a contraction of *ba* ‘to’ + =*ana* (post-position) ‘over, in’; while the absolute preposition *čana* ‘in’ is a contraction of *ja* + =*ana* (see MacKenzie 1966: 55). Another example in the Pāwa material is *pā* ‘to that’ (3), a contraction of *ba* ‘to’ + ā ‘distal demonstrative adjective’.

Another process is assimilation of certain vowels. For example, the final short vowel /a/ in a stem undergoes assimilation to an initial vowel ē in a following suffix (see MacKenzie 1966: 11), as illustrated in Pāwa: *hāna* ‘spring’, but *hān-ēwa* ‘spring-INDE’ [P.M. 2017]. A stem-final vowel ī can be assimilated to a following *a*: *twānī* ‘stone’, but *twān=a* ‘stone=COMP’ (22). The initial vowel *a* of the definiteness marker -*aka* ‘DEF.M’ (-*akē* ‘DEF.F’, -*akē* ‘DEF.PL’) may also be assimilated: *zāwro-ka=t* ‘child-DEF.M=2SG’ (‘your child’) (34).

4. Morphosyntax

4.1. Nominal morphosyntax

4.1.1. Nominal morphology

The main categories associated with Hawrāmī nominals are grammatical gender, case, number, definiteness, and indefiniteness.¹⁹

¹⁹ The categories of gender, case, and number are also marked on adjectives in agreement with nouns they modify (MacKenzie 1966: 13). Due to space limits, this topic is not discussed in our study.

Grammatical gender as masculine or feminine is evident in the form of the noun stem. A masculine noun stem can end in a consonant, a stressed vowel *-a*, *-ī*, *-o*, or *ū*, or often in the stressed vowel *-ā* (MacKenzie 1966: 13). For example (Pāwa): *hē'la* ‘egg’, *han'nī* ‘watermelon’, *d'ro* ‘lie’, *žū'žū* ‘hedgehog’, and *pī'ā* ‘man’ [P.M. 2017]. A feminine noun stem can end in the stressed vowel *-ē*, unstressed *-a* or *ī*, or sometimes stressed *-ā*. For example: *samā'wara* ‘samovar’, *'twanī* ‘stone’, *wā'le* ‘sister’, *'āwī* ‘water’, and *dā'dā* ‘grandmother’ [P.M. 2017].

Case is distinguished as direct and oblique (see MacKenzie 1966: 14–15). Direct case is not overtly expressed on singular nouns. It is fused with number on plural nouns. Oblique case for singular nouns distinguishes masculine and feminine. It is also fused with number on plural nouns (see below). Oblique case on masculine singular nouns is marked with *-ī* (-y following vowel). Oblique case on feminine singular nouns is *-ē* (allomorphs: stem-final unstressed *-a* assimilates to *-ē*; stem-final *-ā* is raised to *-e* [ɛ] in Nawsūd, but to *-āy* in Pāwa). For example: *warg* ‘wolf’ (direct), *warg-ī* ‘wolf-OBL.SG.M’, *mangāw* ‘cow’ (direct), *mangāw-ē* ‘cow-OBL.SG.F’, *karga* ‘hen’ (direct); *karg-ē* ‘hen-OBL.SG.F’, *adā* ‘mother’ (direct) [examples from P.M. 2017], *ad-e* ‘mother-OBL.SG.F’ [D.M. 1966: 15] (Nawsūd, oblique), but *ad-āy* ‘mother-OBL.SG.F’ [P.M. 2017] (Pāwa, oblique).

Oblique case marks nouns functioning as: 1) the complement of an adposition; 2) the dependent element in a genitive (ezāfe) construction; and 3) the argument expressing object in a present tense clause. It can also mark 4) the agent of a past transitive clause (see MacKenzie 1966:13), though there are no instances of this in our Pāwa text. Examples include (Pāwa): 1) *ja warg-ī matarsū* ‘from wolf-OBL.SG.M I fear’ (‘I fear the wolf’) [P.M. 2017]; 2) *dāstāna=w dādā-y* ‘story=ez_2 grandmother-OBL.SG.F’ (‘Grandmother’s story’) (1); 3) *mansūr-ī nimažnāsū* ‘Mansūr-OBL.SG.M I do not know’ (‘I don’t know Mansūr’) [P.M. 2017]; and in Nawsūd, 4) *wārān-ī nāst* ‘rain-OBL.SG.M did not permit’ (‘the rain did not permit’) [D.M. 1966: 51].

Number is marked and also fused with case. Plural number and direct case is expressed by *-ē* (stem-final vowel *ā* to *e* [ɛ], or stem-final vowel *a* to *ē*). Plural number and oblique case is expressed by *-ā* (stem-final *ā* is followed by *-yā*; stem-final *ē* is assimilated to *ā*) (see MacKenzie 1966: 14–15). In Pāwa, the form *-ā* occurs in written texts, whereas *-ān* is found in spoken speech. Examples include: *warg-ē* ‘wolf-DIR.PL’ [P.M. 2017]; *yān-ē* ‘house-DIR.PL’ [D.M. 1966: 14] (stem is *yāna*); *ad-e* ‘mother-DIR.PL’ [D.M. 1966: 15] (stem is *adā*); *warg- ī* ‘wolf-OBL.PL’ [P.M. 2017] (Pāwa, written form); *warg- 'ān* ‘wolf-OBL.PL’ [P.M. 2017] (Pāwa, spoken form); *yān-ā* ‘house-OBL.PL’ [D.M. 1966: 14] (stem is *yāna*); and *adā-yā* ‘mother-OBL.PL’ [D.M. 1966: 15] (stem is *adā*).

Indefiniteness and singularity is marked in Nawsūd on a masculine singular noun with *-ēw* and a feminine singular noun with *-ēwa*. In Pāwa, there is no distinction of gender in this form, and only the form *-ēwa* occurs. Stress remains on the stem of the noun. Examples include: *kur-ēw* ‘son-INDEF.M’ [D.M. 1966: 66]; *warg-*

ēwa ‘wolf-INDF’ [P.M. 2017] (Pāwa); and *yān-ēw* ‘house-INDF.M’ [D.M. 1966: 16] (stem is *yāna*); *yān-ēwa* ‘house-INDF’ [P.M. 2017] (Pāwa). There is another indefinite marker, *-ya*, for example: *dēq-ya* ‘sorrow-INDF_2’ (22).

Nouns marked for plural (with case), but which show no additional marking of definiteness/indefiniteness, can express a sense of indefinite plural, as ‘some’ (see MacKenzie 1966: 15). For example: *yān-ē* ‘house-DIR.PL’ (i. e., ‘some houses’) [P.M. 2017].

Nouns are marked for definiteness, or rather identifiability, by *-a'ka* (masculine nouns) and *-a'kē* (feminine nouns), identical to *-a'kē* (plural) (following certain stem-final vowels, the initial vowel of the morpheme is assimilated). For example: *warg-aka* ‘wolf-DEF.M’ [P.M. 2017]; *yānaka* ‘house.DEF.M’ [D.M. 1966: 16]; *adā-kē* ‘mother-DEF.F’ [D.M. 1966: 16]; *palawar-akē* ‘bird-DEF.PL’ [P.M. 2017]; and *ade-kē* ‘mother-DEF.PL’ [D.M. 1966: 16].²⁰

4.1.2. Demonstrative particle and the demonstrative construction

The demonstrative particle distinguishes gender: *=a* ‘DEM.M’ and *=ē* ‘DEM.F’ (see MacKenzie 1966: 16–17). The particle is stressed and attaches as an enclitic normally to the final element of a noun phrase construction consisting of a noun as head, preceded by the proximal demonstrative adjective *ī* or the distal form *ā*. The gender of the particle agrees with that of the head noun. The construction is exemplified here (Pāwa): *ī hanār=a* ‘PROX.DEM.ADJ pomegranate=DEM.M’ (‘this pomegranate’); and *ā žan=ē* ‘DIST.DEM.ADJ woman=DEM.F’ (‘that woman’) (stem is *žanī*) [P.M. 2017]. This particle can also simply attach to a noun phrase in vocative use, as in Pāwa: *gīyān=a* ‘life/soul=DEM.PTCL.M’ (‘O beloved one’) (21) (see also MacKenzie 1966: 20).

4.1.3. Ezāfe and compound particles in noun phrase constructions

There are two main forms of ezāfe particles (both unstressed). The first is *=ī* (=y following some vowels; absent following *-ī*, stressed *-ī*, or *-ē*), here glossed as ‘EZ_1’. The second is *=ū* (=w following vowel), glossed as ‘EZ_2’. The ezāfe particle attaches to a head noun phrase and links it with a dependent element in what is termed the “ezāfe construction” (see MacKenzie 1966: 17–19). The ezāfe particle *=ī* links a head noun phrase with an attributive adjective (or a noun in apposition), while the ezāfe particle *=ū* links a head noun phrase with another noun phrase, pronoun, preposition, or adverb.

²⁰ The segmentation between the stem-final vowel of *yāna* and the initial vowel of *-aka* is not clear.

Another linking particle is the “compound marker” =*a*. The compound marker has a particular use in definite noun phrases (discussed in relationship to =*ī* ‘EZ_1’ below).

The ezāfe construction with =*ī* is exemplified here: *z(i)msān=ī čapat* ‘winter=EZ_1 dirty’ (‘dirty winter’) (2); *kitēb=ī sīāw* ‘book=EZ_1 black’ (‘black book’) [D.M. 1966: 17]; and *yān-ēw=ī gawra* ‘house-INDF.M=EZ_1 big’ (‘a big house’) [D.M. 1966: 17] (stem is *yāna*).

The ezāfe construction with =*ī* can be found on a head noun that is unmarked or marked for indefiniteness (the ezāfe then follows that marker; see preceding example, *yān-ēw=ī gawra*). When a head noun is specified for definiteness (marked with either a definiteness marker, or modified by a demonstrative), the ezāfe =*ī* is not used. Instead, the compound marker =*a* is employed in what is termed the “open” compound construction (see MacKenzie 1966: 18). Examples include: *kitēb=a sīāw-aka* ‘book=COMP.D black-DEF.M (‘the black book’) [D.M. 1966: 18]; and *ā kināč=a zarīf=ē* ‘DIST.DEM.ADJ girl=COMP.D beautiful=DEM.F’ (‘that beautiful girl’) [D.M. 1966: 18] (stem is *kināčē*).

The ezāfe construction with =*ū* is illustrated here: *yāna=w xudā-y* ‘house=EZ_2 God-OBL.SG.M’ (‘the house of God’) (46). It can also link several elements: *sar=ū bān=ū mizgī=wa* ‘on=EZ_2 roof=EZ_2 mosque=POSTP#_1’ (‘on the roof of the mosque’) (57).

4.2. Pronouns and the reflexive

4.2.1. Independent pronouns

Hawrāmī has independent as well as clitic forms of personal pronouns. The independent forms consist of a set of first and second person pronouns as well as two sets of demonstrative pronouns with third person reference (for tables, see MacKenzie 1966: 24–25, with minor modifications).

The first and second person pronouns distinguish person and number but not case:

Table 4: First and second person pronouns

Person	Singular	Plural
1	<i>min</i> (Pāwa: <i>amin</i>)	<i>ē'ma</i>
2	<i>to</i>	<i>ši'ma</i>

There are two sets of demonstrative pronouns. Those of the first set can be used pronominally in third person reference and also as demonstratives distinguished for distance, as proximal and distal. Forms of the second set are noted as “purely

demonstrative” (MacKenzie 1966: 24). The sets are shown here (some forms are not segmented further):

Table 5: Demonstrative pronouns, Set 1 and Set 2

Set 1			Set 2	
	Proximal	Distal	Proximal	Distal
<i>Singular Direct</i>	M.	<i>ēd̥</i>	<i>ād̥</i>	
	F.	<i>'ēda</i>	<i>'āda</i>	
<i>Singular Oblique</i>	M.	<i>'ēd̥t̥</i>	<i>'ād̥t̥</i>	<i>īna 'ya</i>
	F.	<i>'ēd̥ē</i>	<i>'ād̥ē</i>	<i>īnē</i>
<i>Plural Direct</i>		<i>'ēd̥ē</i>	<i>'ād̥ē</i>	<i>īnē</i>
<i>Plural Oblique</i>		<i>ēd̥i 'šā</i>	<i>ād̥i 'šā</i>	<i>īnī 'šā</i>

4.2.2. Clitic pronouns

Clitic pronouns distinguish person (first, second, third) and number (singular and plural). (Forms in parentheses are allomorphs following vowels.)

Table 6: Clitic pronouns

Person	Singular	Plural
1	=im (=m)	=mā
2	=it (=t), variant: =iₜ	=tā
3	=iš (=š)	=šā

Clitic pronouns are used in functions that typically require oblique case. Most of the functions occur in our glossed text at the end of this study: 1) possessor in a noun phrase (14), (25), (34); 2) complement of adpositional phrase (26), (48); 3) direct object (O) argument of present tense clause (2), (13); 4) agent (A) argument of past tense clause (11), (24), (36), (41), (54); 5) specifying referent on reflexive *wē* (4), (62); 6) experiencer or possessor in a copula clause (5), (34), (51); and 7) indirect object argument (32).

4.2.3. Reflexive

The form of the reflexive is *wē*. It is always followed by a clitic pronoun that specifies the person and number of the referent. It is used to express a referent that is identical to that of an antecedent within the clause as agent or subject (or in an imperative, the addressee, see example in MacKenzie 1966: 27). Uses of the reflex-

ive are (with examples from our glossed text): 1) to express possessor in possessive constructions (53); 2) as complement of an adpositional phrase (43); 3) to express a type of emphasis involving the referent (4); and 4) to express direct object (O) in transitive clauses when the referent is the same as the agent (A), for example, *wē=š fārā* ‘REFL=3SG disguise.PST’ (‘she/he disguised her/himself’) [D.M. 1966: 27].

4.3. Verbal morphosyntax

4.3.1. Stems

A typical verb has two stems: past and present. These form the basis for finite verb constructions. The past stem forms the basis for the non-finite verb constructions of the participle and the “infinitive verbal noun”. The participle is formed from the past stem and the ending -'a (unmarked or masculine), -'ē (feminine), or -'e (plural) (MacKenzie 1966: 36). The infinitive verbal noun is built with the ending -'(a)y (MacKenzie 1966: 28).

The form of a verb stem can be simple or complex. Examples of simple verb stems include: *kar-* ‘do.PRS’ and *kard-* ‘do.PST’; *zān-* ‘know.PRS’ and *zānā-* ‘know.PST’; (*w*)*āč-* ‘say.PRS’ and *wāt-* ‘say.PST’; and *ār* ‘bring.PRS’ and *āwrd* ‘bring.PST’ (see glossed text in Section 5).

Complex forms can be of three types: 1) a “light verb” construction; 2) a “postverbal” construction; and 3) a “preverbal” construction. The light verb construction is formed with a nominal, an adjective, or an adverb as a first element and a “light verb” as the second element. For example: *tamāša kar-* ‘looking (noun) do.PRS’ (‘to look at’) [D.M. 1966: 110]; and *bar ūt-* ‘out (adverb) go.PST_2’ (‘go out’) (9).

The postverbal construction involves a simple stem combined with a particle *an'a*, *ar'a*, or *aw'a* (Pāwa) (Nawsūd: ɔ following ī, ū, or a consonant). These particles also occur in other contexts as postpositions (MacKenzie 1966: 31). They can occur on a verb and give it a different lexical sense. For example: *nīšt-* ‘sit.PST’ > *nīštara-* ‘sit.down.PST’; *kard-* ‘do.PST’ > *kardana-* ‘put.on.PST’; and *ward-* ‘eat.PST’ > *wardēwa-* ‘drink.PST’ [D.M. 1966: 31].

The preverbal construction is built with a simple verb stem and a preverb, *hur* ‘up’ (Pāwa *hor*), which also gives the stem a new sense. For example: *gēr-* ‘keep.PRS’ > *hor gēr* ‘lift up.PRS’ (20).

As noted by MacKenzie (1966: 48–49), the present causative stem is formed by adding *-n* added to a present stem, while the past causative stem is formed by adding *-nā* to a present stem: *giraw-* ‘weep’ > *girawn-* ‘make.weep.PRS’ and *girawnā-* ‘make.weep.PST’ [D.M. 1966: 49]. MacKenzie also notes that the present passive stem is built with the addition of *-ia*, while the past passive stem is formed by adding *-īa* to a present transitive stem, while the past passive stem is formed by adding *-ī'ā-* to a present transitive stem. For example: *wān-* ‘read’ > *wānīa-* ‘be.read.PRS’ and *wānīā-* ‘be.read.PST’ [D.M. 1966: 48].

4.3.2. Prefixes

The following prefixes indicate mood, aspect, and negation (see MacKenzie 1966: 32):

Table 7: Verbal prefixes

Use	Form
Indicative	<i>mi-</i> (<i>Nawsūd</i>) or <i>ma-</i> (<i>Pāwa</i>) (<i>m-</i> preceding vowel or absent on some stems)
Subjunctive, Imperative	<i>bi-</i> (<i>b-</i> preceding vowel; or absence of marking and instead a stress shift to first syllable of stem)
Negation (Present Indicative and Prohibitive constructions)	<i>ma-</i> (<i>nim-</i> preceding <i>ā</i> and <i>a</i>), <i>ni-</i> (<i>Pāwa</i>)
Negation (constructions other than Present Indicative)	<i>na-</i>

4.3.3. Verb suffixes: Present

The following verb suffixes are used in finite constructions formed from the present stem and also in those formed from past stems in transitive clauses. Most language material in the following tables is from MacKenzie (1966: 33), with our glosses and notes. Forms in parentheses are normally the allomorphs following vowels.

The present suffixes are shown here. The form *-ūna* ‘-1SG.PRS’ is found in *Pāwa* (as well as in *Nawsūd* and *Nodša*) as a variant for first person singular. In our glossed text, the form mainly occurs in subjunctive constructions, as in (13). There also appears to be a second singular form *-īna*, as in (24) (not included in the table below; more study is needed). The present suffixes function to mark the person-number of the subject (S) argument of intransitive verbs and agent (A) of transitive verbs in present tense clauses.

Table 8: Verb suffixes: Present

Person	Singular	Plural
1	<i>-ū</i> (- <i>w</i>) <i>-ūna</i>	<i>-mē</i> (- <i>y'mē</i>)
2	<i>-ī</i> (- <i>y</i>)	<i>-dē</i> (- <i>y'dē</i>)
3	<i>-o</i>	<i>-ā</i> (<i>Pāwa</i> : <i>ān</i>)

The Imperfect is unusual in that even though it is formed from a present stem, it is used to denote situations occurring in past time, with imperfective aspect. The

suffixes mark the person-number of the subject (S) and agent (A) argument in Imperfect constructions.

Table 9: Verb suffixes: Imperfect

Person	Singular	Plural
1	- 'ēnē	- 'ēnmē
2	- 'ēnī (Pāwa: -ēṣī)	- 'ēndē
3	- 'ē	- 'ēnē

The suffix in an Imperative construction indicates the addressee:

Table 10: Verb suffixes: Imperative

Person-number	Form
2SG	-a (unmarked following vowel)
2PL	-dē (-ydē)

4.3.4. Verb suffixes: Past

A different set of verb suffixes is found in constructions formed from past stems of intransitive verbs. These suffixes mark the person-number of the subject (S) argument in past intransitive clauses. They also sporadically mark the person-number of the object (O) argument in past transitive clauses. For third person singular, the suffix also distinguishes gender (with masculine as unmarked).

Table 11: Verb suffixes: Past

Person	Singular	Plural
1	-ā(nē)	-īmē (-ymē), (-aymē)
2	-ī (-y), (-ay following ī)	-īdē (-ydē) (-aydē)
3	(Unmarked for masc.) -a (fem.)	-ē (-y)

4.3.5. Copular forms

The next table shows the copular forms (as found in MacKenzie 1966: 34, 37, 38). Present forms attach as enclitics; and the third singular masculine allomorph =n follows a stressed vowel, while the allomorph =na follows an unstressed vowel:

Table 12: Copula verb: Present, Past, and Imperfect forms

Person-number	Present	Past	Imperfect
1SG	=anā (=nā)	'bīā(nē)	'bēnē
2SG	=anī (=nī)	'bīay	'bēnī (Pāwa: bēšī)
3SG MASC	=ā (=n, =na)	bī	bē
3SG FEM	=ana (=na)	'bīa	bē
1PL	=anmē (=nmē)	'bīaymē	'bēnmē
2PL	=andē (=ndē)	'bīaydē	'bēndē
3PL	=anē (=nē)	'bīē	'bēnē

The copula forms the basis for constructions with a noun phrase complement (16), (*čarma* is a feminine noun) probably an adpositional complement (no example available), and an adjectival complement (12), (61). The copula also is used in possessive constructions (34) and experiencer constructions (5), (15).

There is another third person singular form of the copula, *mawo*, only attested once in our text (22), and also a further unanalyzed copula, *nā* (5), (22).

4.3.6. Verb constructions

The finite verb constructions are summarized here:

Table 13: Finite verb constructions

Construction	Description	Example
Present Indicative	<i>mi-</i> (Pāwa: <i>ma-</i>), Present stem, Present verb suffix	<i>ma-wīn-ū</i> ‘I see’ (15)
Present Subjunctive	<i>bi-</i> , Present stem, Present verb suffix	<i>bi-sān-o</i> ‘he/she may buy’ (37)
Present Continuous	<i>ma-</i> , reduplicated stem adverb with -āy, Present stem or Imperfect, Present verb suffix ²¹	<i>rāmāy maram-ūna</i> ‘I am now running’ [M&B 2004]
Present Imperative	<i>bi-</i> , Present stem, Imperative suffix	<i>bi-dā</i> ‘Give!’ [P.M. 2017]
Present Prohibitive	<i>ma-</i> negation, Present stem, Imperative suffix	<i>ma-dā</i> ‘Don’t give!’ [D.M. 1966: 38]

²¹ This construction is noted by MacKenzie (1966: 50) as a “Continuous” tense formed from the stem-based adverb and either the Present Indicative or Imperfect.

Construction	Description	Example
Imperfect	Present stem, Imperfect verb suffix	<i>kar-ēnē</i> ‘I was doing’ (44)
Past Perfective (Simple Past)	Past stem, Past verb suffixes (clitic pronoun with transitive verbs)	<i>lū-wā</i> ‘he went’ (54)
Past Conditional	Past stem, Past Conditional verb suffixes	<i>lū-enē</i> ‘I would have gone’ [D.M. 1966: 46]

The Present and Past Perfect constructions, formed from a combination of finite and non-finite verb forms, are summarized here:²²

Table 14: Perfect constructions

Construction	Description	Example
Present Perfect	Past participle, Present copula	<i>witē=nmē</i> ‘we have slept’ [D.M. 1966: 39]
Past Perfect (Pluperfect)	Past participle, Imperfect copula	<i>witē=bē</i> ‘she had slept’ [D.M. 1966: 40]
Perfect Subjunctive	Past participle, Present Subjunctive copula	<i>wita=bo</i> ‘he may have slept’ [D.M. 1966: 39]
Perfect Conditional	Past participle, Past Conditional copula	<i>wita=bē</i> ‘he would have slept’ [D.M. 1966: 40]

4.3.7. Clause argument indexing and alignment

The patterning of clause argument indexing in finite verb constructions is conditioned by the tense and transitivity of the verb. In a present tense clause, subject (S) of an intransitive verb or agent (A) of a transitive verb will be indexed on the verb by means of a present suffix.

In past tense clauses, S of an intransitive verb is indexed on the verb with a past suffix. O of a transitive verb is indexed with a past suffix. (This analysis is provisional, since we do not have examples of all person-number markers in our text.)

²² Note that the participle in the Present and Past Perfect constructions can be marked for masculine or feminine gender, or for plural number. It is also noteworthy that the copula in the Past Perfect construction is a special form of the Imperfect copula in which gender is distinguished in all singular person forms. In contrast, the independent form of the Imperfect copula does not have gender distinctions (see next table with the Imperfect copula).

In past tense transitive clauses, A is indexed by a clitic pronoun. If A is expressed by a full noun phrase, the clitic pronoun still occurs as resumptive reference.²³

In terms of case and alignment, a noun phrase is marked with oblique case when expressing O in a present tense transitive clause, or A in a past tense transitive clause. (We do not find an oblique-marked noun phrase expressing A in our text, but an instance is found in MacKenzie 1966: 51, *wārān-ī n-āst* ‘rain-OBL.SG.M NEG_2-let.PST’ ‘the rain did not permit.’)

Types of patterning are exemplified by sentences from our glossed text:

Present tense clauses:

S in present tense clause with present suffix: (7 (*magnū*)), (64).

A in present tense clause with present suffix: (4), (8), (21), (40).

Past tense clauses:

S in past tense clause with past verb suffix: (52).

O in past tense clause with past suffix: (36), (41 (*kard-a*)), (49 (*kard-a, hormižā-y*)). It is also observed in (39) that the past suffix *-ānē* ‘1SG.PST’ indexes an indirect object, not O.

A in past tense clause with clitic pronoun: (20), (41 (=*mān*, =*m*)), (43).

4.4. Other observations

There are other word classes not discussed here, including the class of absolute prepositions (such as *čanī* ‘with’, *pora* ‘on, to’, *pana* ‘to’). Further areas of morphosyntax that deserve more study include the use of the preposition *pay* ‘to’ and the directional particle *=a*; the use of the additive particle *=īč* (variant: *=īš*) in discourse; and relativization, including the irregular use of *ka* as a complementizer. These and many other topics remain to be described elsewhere.

5. Sample text in Hawrāmī (Pāwa)

This text is a true story, written in August 2011 by Jahangir Mahmoudveysi. He is from Pāwa and has authored other short stories and poetry in Hawrāmī. He is fifty-six years old. He also read this story aloud as an oral version, which was recorded and transcribed by Parvin Mahmoudveysi in the summer of 2017.²⁴

²³ See MacKenzie (1966: 51).

²⁴ For helpful comments on several aspects of our translation and morphemic glossing of the text, we would like to thank the participants of the Hawrāmī course taught by Parvin Mahmoudveysi at the 2017 Kurdish Language Autumn School at the University of Bamberg, Germany.

Grandmother's Story, by Jahangir Mahmoudveysi

- (1) *dāstāna=w dādā-y*
 story=EZ_2 grandmother-OBL.SG.F
 Grandmother's story.²⁵
- (2) *dādā: xudā n-ār-o=t*
 grandmother: God NEG_2-bring.PRS-3SG.PRS=2SG
ay z(i)msān=ī čapat=u bē kałk
 O winter=EZ_1 dirty=and without benefit
 Grandmother: "May God not bring you (here) (i. e., you are not welcome),
 O dirty and useless winter!"²⁶
- (3) *dādā čēš m-āč-ī pā*
 grandmother what IND-say.PRS-2SG.PRS with.DIST.DEM.ADJ
giz(ī)=ya=w das-ī=t=a ā
 broom=DEM.PTCL.M=EZ_2 hand-OBL.SG.M=2SG=DEM.PTCL.M DIST.DEM.ADJ
warw-ē p(ay)čī ma-pižgn-ī
 snow-OBL.SG.F/DEM.PTCL.F why IND-scatter.PRS-2SG.PRS
 (Āzād speaks) "Grandmother, What are you saying? Why are you scattering
 that snow with that broom in (literally, of) your hand (literally, that hand of
 yours)?"²⁷
- (4) *řoła āzād gīyān wāz=im čana b-ār-a*
 dear_child Āzād life/soul open=1SG in SBJV-bring.PRS-2SG.IMP
wē=m ma-zān-ū čēš ma-kar-ū
 REFL=1SG IND-know.PRS-1SG.PRS what IND-do.PRS-1SG.PRS
 (Grandmother speaks) "My dear beloved child Āzād, leave me in peace.
 I myself know what I'm doing.

²⁵ The form of the feminine oblique -y is an allomorph of -āy, as found in Pāwa Hawrāmī; Nawsūd is -e.

²⁶ The word *xudā* follows Pāwa pronunciation. The speaker originally pronounced it as *xwā* here and elsewhere.

²⁷ The gloss as 'OBL.SG.F/DEM.PTCL.F' is provisional. Both endings would be expected in this context, but only one form appears. The oblique is a suffix and the demonstrative particle is a clitic, but otherwise these two endings are phonologically identical. A similar issue is found below in (7).

- (5) *amin qīnī=m=ana* *z(i)msān-ī*
 1SG hate=1SG=COP.PRS.3SG.F winter-OBL.SG.M
z(i)msān pēsa tīr-ya nā gīyān=m=ana
 winter like arrow-INDEF_2 is.NA life/soul=1SG=POSTP#_2
 I hate winter. Winter is like an arrow in my soul.
- (6) *ja ūranga=ū čarm-ē waš=im*
 from color=EZ_2 white-OBL.SG.F pleasure=1SG
ni-m-ay
 NEG_1-IND-come.PRS.3SG
 I don't like the white color (literally, color of white).²⁸
- (7) *řanga=ū čarm-ē ma-wīn-ū ma-gn-ū*
 color=EZ_2 white-OBL.SG.F IND-see.PRS-1SG.PRS IND-fall.PRS-1SG.PRS
yād=ū ā hatīm-ya
 memory=EZ_2 DIST.DEM.ADJ irresponsible_guy-OBL.SG.M/DEM.PTCL.M
 (When) I see the white color, I remember that irresponsible guy.
- (8) *ba hāsān-ī līwā=u na-tāwā=š*
 with ease-OBL.SG.M go.PST_1=and NEG_2-can.PST=3SG
hīč=īš kar-o
 nothing=ADD_2 do.PRS-3SG.PRS
 He went (around) with ease, and he could not do anything either.
- (9) *pēsa jarda-k-ān bar-šī=u*
 like thief-DEF.M-PL out-go.PST_2=and
ā k(i)nāč=a xās-ē=m=a
 DIST.DEM.ADJ daughter=COMP.DEM good-DEM.PTCL.F=1SG=DEM.PTCL.M
ja das dā
 from hand give.PST
 Like the thieves, he fled, and I lost that good daughter of mine.²⁹

²⁸ The word *čarma* 'white (color)' is a feminine noun.

²⁹ The form of the demonstrative particle as =a (normally masculine) can agree with a head noun of either gender when other morphemes occur between the head noun and the particle (see MacKenzie 1966: 16).

- (10) *ay ūrla gīyān āzīz-akē=m duwagīyān-akē*
 O dear_child life/soul dear_one-DEF.F=1SG pregnant-DEF.F
kē haq=it bi-sān-o ay xuđā
 who right=2SG SBJV-take.PRS-3SG.PRS O God
īna=t=a čanī qabūl kard
 PROX=2SG=DEM.PTCL.M how accept do.PST
 O dear beloved child, my dear one, you were pregnant (literally, the pregnant one), who may take your right? O God! How did you accept this?
- (11) *pī hatīm-ān=a pēsa=t=a*
 to.PROX.DEM.ADJ orphan-PL=DEM.PTCL.M such=2SG=DEM.PTCL.M
p(ay)čīt kard
 why do.PST
 Why did you do such (as this) to these orphans?"
- (12) *dāđā gīyān p(ay)čīt dēwyāy=nī*
 grandmother life/soul why gone_mad=COP.PRS.2SG
 (Āzād speaks) "Dear Grandmother, why have you gone mad (i. e., are distraught)?"
- (13) *āzāđ gīyān b-āz-a=m bā hāwār*
 Āzād life/soul SBJV-let.PRS-2SG.IMP=1SG let_it_be shout_for_help
kar-ūna bā ī warw-ān=a
 do.PRS-1SG.PRS.SBJV let_it_be PROX.DEM.ADJ snow-PL=DEM.PTCL.M
šēwn-ūna
 shovel.CAUSE.PRS-1SG.PRS.SBJV
 (Grandmother speaks) "Dear Azad, let me, let me shout, let me shovel (away) this snow (literally, these snows)!"
- (14) *dāđā gīyān wasē=n lic-akē=t*
 grandmother life/soul enough=COP.PRS.3SG.M lip-DEF.PL=2SG
zard-ē bīyē=nē=u ī rāngā=t
 yellow-PL.AGR COP.PTCP.PL=COP.PRS.3PL=and color=2SG
wāryā=yāna
 change.PTCP=COP.PRS.3SG.F
 (Āzād speaks) "Dear Grandmother, it is enough, your lips have become pale (literally, yellow) and your color (i. e., of your face) has changed."³⁰

³⁰ This realization of the second person singular clitic pronoun as =*t* is a variant of =*t*, allomorph of =*it* (see also 20, 37). The different realization as =*it* in the subsequent instance could be due to the influence of other languages.

- (15) *řołta gīyān amin qīna=m=ana řang(a)=ū*
 dear_child life/soul 1SG hate=1SG=COP.PRS.3SG.F color=EZ_2
čarm-ē waxt-ēwa řang(a)=ū čarm-ē
 white-OBL.SG.F time-INDF_1 color=EZ_2 white-OBL.SG.F
ma-wīn-ū ma-dēwy-aw
 IND-see.PRS-1SG.PRS IND-go_mad.PRS-1SG.PRS
 (Grandmother speaks) “Dear beloved child, I hate white color. When I see
 white color, I go mad.
- (16) *ā šaw=a g(i)rd=ū dagā-y čarma*
 DIST.DEM.ADJ night=DEM.PTCL.M all=EZ_2 village-OBL.SG.M white
bīya=bē kip=u
 COP.PTCP=COP.IPF.3SG closed_up=and
bē dang sirtēwa n-ē bar
 without sound/voice whisper.INDF.F NEG_2-come.IPF.3SG out
 That night, all of the village had become white; completely silent and
 without sound, (even) a whisper was not coming out (of it).
- (17) *warwa pēsa lēf-ī kēšyā=bē*
 snow like blanket-OBL.SG.M pulled_over.PTCP=COP.IPF.3SG
sar=ū dagā-y=ra
 head_2=EZ_2 village-OBL.SG.M=POSTP#_3
 Snow like a blanket was pulled over the village.³¹
- (18) *kas zāt=iš nawē*
 someone/no-one daring_courage=3SG NEG.COP.PST.3SG
ja yāna=na b-ay bar
 from house=POSTP#_2 SBJV-come.PRS.3SG out
 No-one had (the) courage to come out of the house.

³¹ The participle in the Past Perfect in Pāwa (here: *kēšyā=bē*) may be marked in certain instances for gender (masculine/unmarked or feminine) or for plural number. In most of the text occurrences here and elsewhere, the marking is assumed to be assimilated to the final -ā vowel or it may be identical as -a to an unmarked form, and it is thus left unglossed.

- (19) *ā basazwān=ē ba lama=y*
 DIST.DEM.ADJ poor_creature=DEM.PTCL.F with belly=EZ_1
pař-a pā k(i)rēw=ana čanī=w
 full-F.AGR at.DIST.DEM.ADJ snowstorm=POSTP#_2 with=EZ_2
šū-aka=yš mašyāya bi-l̄wā-ya
 husband-DEF.M=3SG should.PST SBJV-go.PST_1-NA
 That poor one (who was) pregnant, (that) she should go (out) in that
 snowstorm with her husband!”³²
- (20) *dādā das hor-gēr-a warwa-kē=t*
 grandmother hand up-keep/lift.PRS-2SG.IMP snow-DEF.F=2SG
tāwnā=wa ītr wasē=n
 melt_cause.PST=POSTP#_1 anymore enough=COP.PRS.3SG.M
 (Āzād speaks) “O Grandmother, let it be, you melted the snow. It is enough
 now.”
- (21) *gīyān=a to čēš ma-zān-ī*
 life/soul=DEM.PTCL.M 2SG what IND-know.PRS-2SG.PRS
 (Grandmother speaks) “O beloved one, what do you know?
- (22) *dēq-ya nā sar=ū dit-ī=m=ana*
 sorrow-INDF_2 is.NA head_2=EZ_2 heart-OBL.SG.M=1SG=POSTP#_2
enna=w twan=a gawra-kē=w arab-ak-ān mawo
 so_much=EZ_2 stone=COMP big-DEF.F=EZ_2 Arab-DEF.M-PL be.3SG
 There is a feeling of sorrow over my heart, it is (as big as) the big stone of
 the Arabs.³³

³² The idiom *basazwān*, literally ‘bound/closed tongue’, is normally used to refer to an animal, which does not speak, but it can also refer to a person who is quiet and without fault.

³³ The ‘big stone of the Arabs’ refers to the stone in a ritual practiced by pilgrims to Mecca. In the ritual, pilgrims throw pebbles at the stone, which is believed to represent the devil.

- (23) *ā bābahīz=a ghazab-bag-ī*
 DIST.DEM.ADJ lecher/[expletive]=DEM.PTCL.M rage-Bag-OBL.SG.M
m-āč-ūna zor=iš pay šū-waka=yš
 IND-say.PRS-1SG.PRS.SBJV pressure=3SG to husband-DEF.M=3SG
āʷrda=bē
 bring.PTCP=COP.IPF.3SG
 That lecher (i. e., [expletive]), I mean Honorless Bag, had put (literally,
 brought) pressure on her husband.³⁴
- (24) *wāta=bē=iš*
 say.PTCP=COP.IPF.3SG=3SG
mašyo t(i)fang-ēwa=m pay b-ār-īna
 should.PRS gun-INDEF_1=1SG for SBJV-bring.PRS-2SG.PRS.SBJV
 He had said: "You should bring me a gun."
- (25) *wa(r)na sawāy pīyā-k-ān=im ma-kyān-ū*
 if_not tomorrow man-DEF.M-PL=1SG IND-send.PRS-1SG.PRS
gīyān=(i)t
 life/soul=2SG
 If not, tomorrow I will send my men for your life (i. e., to take your life)."
- (26) *ād=īč ſon-a=š gim-a bo*
 DIST.M_1=ADD_1 trace-F.AGR=3SG lost-F.AGR SBJV.COP.PRS.3SG
zor=iš na-mařyā
 power=3SG NEG_2-break.PST
mil=š=ara r̄ā=w bar-šīya=yš girt=a
 on=3SG=POSTP#_3 way=EZ_2 out-go.INF_2=3SG take.PST=DIR.PTCL
war
ahead
 As for him, may (every) trace (literally, the trace) of him be lost, he had
 no power, going out, he fled.
- (27) *bafrāw=iš dā war*
 Bafrāw=3SG give.PST ahead
 He had Bafrāw go ahead (literally, he put Bafrāw ahead).

³⁴ The word *ghazab*, literally, 'rage', is used here with *bag* as a proper name to express the referent's lack of honor and fairness. A figure from earlier times, the *bag* of a village had a role similar to that of a mayor. A *bag* was often unfair in dealings with the village people.

- (28) *pā nīmašaw=a dilē=w ā*
 at.DIST.DEM.ADJ midnight=DEM.PTCP.M in_middle=EZ_2 DIST.DEM.ADJ
k(i)řēwa-y=na ja dagā=na
 snowstorm-OBL.SG.M=POSTP#_2 from village=POSTP#_2
bar-a-šī-ya
 outside-F.AGR-go.PST_2-3SG.F
 At (literally, at that) midnight in the middle of that snowstorm, she went out of the village.
- (29) *am(i)n=īš kor-a bīyē=bēnē*
 1SG=ADD_2 blind-F.AGR COP.PTCP.F=COP.IPF.1SG.F
hīč=im pay na-k(i)ry-ē
 nothing=1SG for NEG_2-be_done.PRS-3SG.IPF
 As for me (i.e., Grandmother), I had become blind, it was not possible for me to do anything.
- (30) *zāt=īč=im nawē hīč*
 daring_courage=ADD_1=1SG NEG.COP.PST.3SG nothing
b-āč-ūna ā
 SBJV-say.PRS-1SG.PRS.SBJV DIST.DEM.ADJ
gamālbačaba=ya ažnās-ēnē
 dog_father[expletive]=DEM.PTCP.M know.PRS-1SG.IPF
 I did not have courage to say anything, I knew that rotten guy (literally, [expletive]).
- (31) *zot=u jargsaxt=u bē-řahim bē*
 bastard=and stubborn=and merciless COP.IPF.3SG
 He was a bastard and stubborn and merciless.
- (32) *ħama gētnā=š=o čēš=šān*
 Hama report.PST=3SG=POSTP#_1 what=3PL
sara-āma=n
 head_1-come.PTCP.M=COP.PRS.3SG.M
 Hama reported what had (literally, has) happened to them.³⁵
- (33) *ay řoṭa gīyān řoṭa=y*
 O dear_child life/soul dear_child=EZ_1
kizol-a=m hatīm=a jwān=a-marg-akē
 innocent_poor_child-F.AGR=1SG orphan=COMP.D young=COMP.D-dead-DEF.F
 O dear beloved child, my poor, innocent dear child, dead-too-young orphan!

³⁵ The Present Perfect is used in reported speech introduced with a verb in Past Indicative.

- (34) *zāwṛo-ka=t či gunāḥ-ēwa=š bē*
 child-DEF.M=2SG what sin-INDF_1=3SG COP.IPF.3SG
 Your child, what sin did she have?
- (35) *hatā kay ĥazrawa=w to kēš-ūna*
 until when suffering=EZ_2 2SG pull/endure.PST-1SG.PRS.SBJV
 How long should I endure your suffering?
- (36) *čwār zāwṛo=ym ba hatīm-ī*
 four child=1SG with orphan-OBL.SG.M
čī ūrō-xrāwy=a=na way
 in.PROX.DEM.ADJ day-bad=DEM.PTCL.M=POSTP#_2 care_for
kard-ē
 do.PST-3PL.PST
 I parented four children without their father in these bad days.
- (37) *kē tāwān=i₄ bi-sān-o*
 who penalty/compensation=2SG SBJV-buy.PRS-3SG.PRS
 Who can compensate you (for this loss)?
- (38) *ay xuđā xās=tān pay na-kar-o*
 O God goodness=2PL to NEG_2-do.PRS-3SG.PRS
 O may God not have mercy on you!
- (39) *āzād gīyān ḥamad pana=š wāt-ānē*
 Azad life/soul Hamad to=3SG say.PST-1SG.PST
 Dear Azad, Hama told me:³⁶
- (40) *kiřewa n-āz-ē čam-ē kar-mē=wa*
 snowstorm NEG_2-let.PRS-3SG.IPF eye-PL.AGR do.PRS-1PL.PRS=PTCL#
 “The snowstorm was not allowing us to open (our) eyes.
- (41) *rā-kē=mān gim-a kard-a hatā lā=m*
 way-DEF.F=1PL lost-F.AGR do.PST-3SG.PST.F until side=1SG
kard=o dī=m bafrāw-a čarm-ē
 do.PST=POSTP#_1 see.PST=1SG Bafrāw-F.AGR white-OBL.SG.F
bīyē=na=wa
 COP.PTCP.F=COP.PRS.3SG.F=PTCL#
 We lost the way, as soon as I turned back, I saw (that) Bafrāw had (literally, has) become pale (literally, white).

³⁶ The Kurdish form of this name is *Hama*. The speaker sometimes uses this name but here he uses *Hamad*.

- (42) *ba ḥāl dang=iš bar ē*
 with situation voice=3SG out come.IPF.3SG
 She spoke very poorly (literally, with situation).
- (43) *časpinā=m wē=m=ara hatā garm=iš*
hold_close.PST=1SG REFL=1SG=POSTP#_3 until warm=3SG
bo=wa
 SBJV.COP.PRS.3SG=PTCL#
 I held her close to myself in order to warm her.
- (44) *bē dasalāt hāwār kar-ēnē kas*
 without power shout_for_help do.PRS-1SG.IPF no_one
diyār=iš nawē
 apparent=3SG NEG.COP.PST.3SG
 I was shouting for help, powerless, no-one was to be found.
- (45) *batām āda hatā ē har sard-a*
 but DIST.F_1 until come.IPF.3SG gradually cold-F.AGR
bē=wa
 COP.IPF.3SG=PTCL#
 But gradually she became cold.
- (46) *w(i)nī=š hatā yāna=w xuđā-y*
 blood=3SG until house=EZ_2 God-OBL.SG.M
mijyā=bē
 pour_out.PTCP=COP.IPF.3SG
 Her blood had poured out endlessly (literally, to the house of God).
- (47) *warwa pēsa xāmēwa čarm-ē*
 snow like cream.INDF white-OBL.SG.F
kēšyā=bē gtyān=(i)š=ara
 pulled_over.PTCP=COP.IPF.3SG body=3SG=POSTP#_3
 Snow, like white cream (literally, a white cream), was spread over her body.
- (48) *tanyā=u bē dasalāt dinyā mijyā=bē*
 alone=and without power world pour_out.PTCP=COP.IPF.3SG
mil=m=ara
 on=1SG=POSTP#_3
 Alone and powerless, the world had fallen apart (literally, poured out) on me.

- (49) *hatā tāwā=m bāwaš-a=m pora kard-a*
 until can.PST=1SG embrace-F.AGR=1SG to do.PST-3SG.PST.F
dimāyīn hanāsa-kē=š=im hor-mižā-y
 last breath-DEF.PL=3SG=1SG up-breathe.PST-3PL.PST
 As much as I could, I embraced her, I breathed her last breaths.
- (50) *gīyān=im larz-ē na-zānā=m*
 body=1SG shake.PRS-3SG.IPF NEG_2-know.PST=1SG
čanī āʷrdē=na=m wārī
 how bring.PTCP.F=COP.PRS.3SG.F=1SG downwards
 My body was shaking. I did not know how I (literally, have) brought her down.
- (51) *ni-ma-zā-ū xatā=m nawyēna*
 NEG_1-IND-know.PRS-1SG.PRS guilt=1SG NEG.COP.PRS.PRF.3SG.F
 I do not know, I did not have any guilt.³⁷
- (52) *tarsa=ū ā pīyākuš(ī)=ya bar-šī-yānē*
 fear=EZ_2 DIST.DEM.ADJ murderer=DEM.PTCL.M out-go.PST_2-1SG.PST
 I fled because of fear of that murderer.
- (53) *žan-kē=w zāwřo-ka=m ja das dā*
 wife-DEF.F=and child-DEF.M=1SG from hand give.PST
šarmanda=w wē=m=u
 shame/mortification=EZ_2 REFL=1SG=and
zāwřo-ka-y=m=anān
 child-DEF.M-OBL.SG.M=1SG=COP.PRS.1SG
 I lost my wife and child. I am full of shame for myself and my child."
- (54) *xwāḥafīzī=š kard hatā āro*
 farewell=3SG do.PST until today
na-dīya=n=(i)m līwā hatā yāwā ērāq
 NEG_2-see.PTCP.M=COP.PRS.3SG.M=1SG go.PST_1 until reach.PST Iraq
 He said (literally, did) farewell, and (since then) until today, I have not seen him; he went until he reached Iraq.
- (55) *řoła gīyān laš=ū bafrāw=ū am(i)n=īš*
 dear_child life/soul body=EZ_2 Bafrāw=EZ_2 1SG=ADD_2
gēlā=wa pay dagā-y
 return.PST=POSTP#_1 to village-OBL.SG.M
 Dear beloved child, as for the body of my Bafrāw, it was returned (literally, it came back) to the village.

³⁷ The form *zā* ‘know.PRS’ here is a variant of the verb *zān*.

- (56) *dagā žitžyā=bē*
 village crowd.PTCP=COP.IPF.3SG
 The village was crowded.
- (57) *sar=ū bān=ū mizgīt=wa kābrā=y dang-zil*
 head_2=EZ_2 roof=EZ_2 mosque=POSTP#_1 guy=EZ_1 loud_voiced
qērnā=š b-a-ydē dam=mān=o
 bellow.PST=3SG SBJV-come.PRS-2PL.IMP mouth=1PL=POSTP#_1
 On the roof of the mosque, a loud-voiced guy bellowed: “Come help us!”
- (58) *kafan=īš wīr=tān na-š-o*
 white_burial_shroud=ADD_2 memory=2PL NEG_2-go.PRS_2-3SG.PRS
 Don’t forget the white burial shroud, too!”
- (59) *duwa kuř-akē=m ī kināč-akē=m*
 two.COMPD son-DEF.PL=1SG PROX.DEM.ADJ daughter-DEF.F=1SG
qīža=w girawa=šān yāwā āsmān
 wailing=and crying=3PL reach.PST heaven
 Both my sons, this (other) daughter of mine, their wailing and crying reached heaven.
- (60) *qērnā=m kafan=iš garak niyan*
 bellow.PST=1SG white_burial_shroud=3SG need NEG.COP.3SG
 I bellowed: “She doesn’t need a burial shroud!”
- (61) *řasūl=īč gētnā=š=o ā řo*
 Rasūl=ADD_1 report.PST=3SG=POSTP#_1 DIST.DEM.ADJ day
yāna=w bag-ī pař bīya=bē ja
 house=EZ_2 Bag-OBL.SG.M full COP.PTCP=COP.IPF.3SG from
mēmān-ī hatā wēraga wašī=šān karda=bē
 guest-OBL.SG.M until evening enjoyment=3PL do.PTCP.M=COP.IPF.3SG
 Rasūl also reported (that), (on) that day, Bag’s house was full of guests, and until evening, they were enjoying their time.
- (62) *tozqāl-ya=yč wē=š nařanjnā=bē*
 tiny_amount-INDF_2=ADD_1 REFL=3SG not_regret.PTCP=COP.IPF.3SG
 He (i. e., Bag) himself was not even a tiny bit sorry.
- (63) *n-āsa=bē=š wašy-aka=šān*
 NEG_2-let.PTCP=COP.IPF.3SG=3SG enjoyment-DEF.M=3PL
bi-šēwy-o
 SBJV-destroy.PRS-3SG.PRS
 He did not let it ruin their enjoyable (i. e., pleasant) time.”

- (64) *dādā wasē=n ā*
 grandmother enough=COP.PRS.3SG.M DIST.DEM.ADJ
bēlēm=a pana ba warw-ē čarm-ē
 shovel=DEM.PTCL.M to to/by snow-PL.AGR white-PL.AGR
čī yān=ana bi-man-o
 in.PROX.DEM.ADJ house=POSTP#_2 SBJV-remain.PRS-3SG.PRS
 (Āzād speaks) “Grandmother, it is enough. Give me the shovel. I should
 not let any white snow remain in this house.”
 [Speaker says: “Jahangir Mahmoudveysi, 20. 05. 1390” (i. e., 11 August
 2011).]

Abbreviations for interlinear morphemic glosses³⁸

Gloss	Definition	Gloss	Definition
_1	Variant 1 (2, 3, etc.)	IND	Indicative
1PL	First person plural	INDEF	Indefinite
1SG	First person singular	INF	Infinitive
2PL	Second person plural	IPF	Imperfect
2SG	Second person singular	M	Masculine
3PL	Third person plural	NA	Not analyzed
3SG	Third person singular	NEG	Negation
ADD	Additive particle	OBL	Oblique
ADJ	Adjective	PL	Plural
AGR	Agreement	POSTP	Postposition
COMPD	Compound marker	PRF	Perfect
COP	Copula	PROX	Proximal
DEF	Definite	PRS	Present
DEM	Demonstrative	PST	Past
DIR	Directional	PTCL	Particle
DIST	Distal	PTCP	Participle
EZ	Ezāfe	REFL	Reflexive
F	Feminine	SBJV	Subjunctive
IMP	Imperative	SG	Singular

³⁸ Most of these abbreviations and symbols follow the Leipzig Glossing Rules (revised version 2008, updated May 2015).

Symbols

- = (Equals sign) indicates an enclitic boundary in examples and glosses.
- (Hyphen) separates the segmentable morphemes in examples and glosses.
- . (Period) separates multiple metalanguage elements that correspond to a single object language element.
- _ (Underscore) separates multiple object language elements that correspond to a single metalanguage element or to a unity of multiple metalanguage elements.
- # (Number sign) refers to a form glossed by word class category.

References

- Benedictsen, Åge Meyer & Arthur Christensen. 1921. *Les dialectes d'Awromān et de Pāwā, textes recueillis par Åge Meyer Benedictsen, revus et publiés avec des notes et une esquisse de grammaire par Arthur Christensen*. København: Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser. VI, 2.
- Bistoon, Abbasi, Amer Gheity & Kerstin Kazzazi. 2013. Semantics of Hawrami kinship terms. *The International Journal of Humanities of the Islamic Republic of Iran* 20(4). 1–21.
- Blau, Joyce. 1989. Gurānî et zâzâ. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 336–340. Wiesbaden: Reichert.
- Christensen, Arthur. 1936. Some new Awromān material prepared from the collections of Åge Meyer Benedictsen. *Bulletin of the School of Oriental Studies* 8(2/3). 467–476.
- Hawrāmānī, Muhamad Amīn. 2001. *Mēzhūy Hawrāmānī*, (617–619). Tehran: Balx. <http://www.hawraman.org/sites/default/files/article/%D8%AA%D9%82%D8%B3%D9%8A%D9%85%D8%A7%D8%AA%D9%8A%D8%A7%D8%B3%D9%8A%D9%85%D8%A7%D9%86.pdf> (accessed 04 October 2017).
- Holmberg, Anders & David Odden. 2008. The noun phrase in Hawrami. In Simin Karimi, Vida Samiian & Donald Stilo (eds.), *Aspects of Iranian linguistics*, 129–151. Newcastle: Cambridge Scholars Publishing. <http://www.ling.ohio-state.edu/~odd़en/The%20Noun%20Phrase%20in%20Hawrami.pdf> (accessed 04 October 2017).
- International Phonetics Association. 2015/2005. *Reproduction of the International Phonetic Alphabet*. <https://www.internationalphoneticassociation.org/content/full-ipa-chart> (accessed 07 October 2017).
- Leezenberg, Michiel. 1993. Gorani influence on Central Kurdish: Substratum or prestige borrowing? *ILLC Prepublication Series*, X-93-03. Amsterdam: Institute for Logic, Language & Computation (ILLC), University of Amsterdam.
- Leipzig Glossing Rules. (Revised version of February 2008, updated May 2015). <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf> (accessed 04 October 2017).
- MacKenzie, David Neil. 1961. The origins of Kurdish. *Transactions of the Philological Society* 1961. 68–86.

- (D.M. 1966=) MacKenzie, David Neil. 1966. *The dialect of Awroman (Hawrāmān-ī Luhōn). Grammatical sketch, texts, and vocabulary*. København: Det Kongelige Danske Videnskabernes Selskab.
- MacKenzie, David Neil. 1987a. Avroman. *Encyclopaedia Iranica* 3.1, 110–111. New York: Center for Iranian Studies, Columbia University.
<http://www.iranicaonline.org/articles/avroman-hawraman-persian-owraman-a-mountainous-region-on-the-western-frontier-of-persian-kurdistan> (accessed 03 January 2015).
- MacKenzie, David Neil. 1987b. Avromani. *Encyclopaedia Iranica* 3.1, 111–112. New York: Center for Iranian Studies, Columbia University.
<http://www.iranicaonline.org/articles/avromani-the-dialect-of-avroman-q> (accessed 05 December 2014).
- MacKenzie, David Neil. 2002. Gurāni. *Encyclopaedia Iranica* 6.4, 401–403. New York: Center for Iranian Studies, Columbia University.
<http://www.iranicaonline.org/articles/gurani> (accessed 14 March 2012).
- Mahmoudveysi, Parvin. 2016. *The meter and the literary language of Gūrānī poetry*. Hamburg: Universität Hamburg dissertation. <http://ediss.sub.uni-hamburg.de/volltexte/2016/7935/pdf/Dissertation.pdf> (accessed 22 June 2017).
- (P.M. 2017=) Mahmoudveysi, Parvin. 2017. Unpublished notes for Hawrāmī (Pāwa).
- (M&B 2004=) Mahmoudveysi, Parvin & Denise Bailey. 2004. Unpublished field notes: Word list and grammar notes for Hawrāmī (Pāwa).
- Mahmoudveysi, Parvin & Denise Bailey. 2013. *The Gorani language of Zarda, a village of West Iran*. Wiesbaden: Reichert.
- Mahmoudveysi, Parvin, Denise Bailey, Ludwig Paul & Geoffrey Haig. 2012. *The Gorani language of Gawrajū, a village of West Iran*. Wiesbaden: Reichert.
- Mahmudweyssi, Parwin [=Mahmoudveysi, Parvin] & Geoffrey Haig. 2009. The typology of modality in some West Iranian languages. In Christine Allison, Anke Joisten-Pruschke & Antje Wendtland (eds.), *From Daēnā to Dīn: Religion, Kultur und Sprache in der iranischen Welt. Festschrift für Philip Kreyenbroek zum 60. Geburtstag*, 41–52. Wiesbaden: Harrassowitz.
- Mann, Oskar & Karl Hadank. 1930. *Mundarten der Gūrān, besonders das Kāndūlāī, Auarāmānī und Bādschälānī, bearbeitet von Karl Hadank*. Berlin: Verlag der Preussischen Akademie der Wissenschaften in Kommission bei Walter de Gruyter & Co.
- Naqšbandī, Šahrām. 1996. *Nezām-e āvāyī guyeš-e Hawrāmī (guneye Pāveh) az dīdgāh-e vāj-šenāsī-ye zāyešī va vāj-šenāsī-ye jozg-e mostaqel [The phonology of Hawrāmī (Pāveh variety), a Generative and Autosegmental perspective]*. Tehran: Allameh Tabatabā’l University, unpublished MA thesis.
- Paul, Ludwig. 2007. Zur Lage der Gōrānī-Dialekte im Iran und ihrer Erforschung. In Maria Macuch, Mauro Maggi & Werner Sundermann (eds.), *Iranian languages and texts from Iran and Turan. Ronald E. Emmerick Memorial Volume*, 285–296. Wiesbaden: Harrassowitz.
- Pāveh Press. 2013. Bištarin mizān-e bāsawādī dar sath-e ostan, marbut be šahrestān-e Pāveh [The highest rate of literacy in the province belongs to the county of Pāveh]. Online article, 29 December, 2013. <http://www.pavehpress.ir/> (accessed 10 August 2017).
- Sadeqi, Amir. No date. *Fasl-e avval: moqeṣiat va taqsimāt-e siyasi va moṣarefi-ye baxš-e Horāmān* [First capital: The situation and political division and introduction to part of Horaman]. Website of Sāzmān-e Mirās-e Farhangi, sanāyeš-e dasti u gardešgari [Organization of Cultural Heritage, Handiworks, and Tourism].

- <http://www.hawraman.org/sites/default/files/article/%D8%AA%D9%82%D8%B3%D9%8A%D9%85%D8%A7%D8%AA%20%D8%B3%D9%8A%D8%A7%D8%B3%D9%8A%20%D9%87%D9%88%D8%B1%D8%A7%D9%85%D8%A7%D9%86.pdf>
(accessed 10 October 2017).
- Shahbazi, Shapur. No date. Iranian geography: Pâveh, the city of a thousand mausoleums. The Circle of Ancient Iranian Studies (CAIS). <http://www.cais-soas.com/CAIS/Geography/paveh.htm> (accessed 28 June 2017).
- Zolfaqari, Ardashir. 2010. The Avromani-speaking area in Iran. *Iran and the Caucasus* 14. 323–330.

4.6. Persian

Ludwig Paul¹

1. General remarks

1.1. Current size and status

Persian enjoys the status of an official or state language in the three Iranophone countries of Iran, Afghanistan, and Tajikistan, where it is called *Fārsī*, *Dari*, and *Tojiki* respectively. It is spoken as a native language by the majority of the population in Iran and Tajikistan, and by a large minority in Afghanistan (whose members are also called ethnic “Tajiks”). In all three countries, it is widely used as a language of communication, education and commerce also by non-native speakers. In Uzbekistan, a vibrant minority of indigenous Tajik speakers still live in the cities of Samarkand and Bukhara. Besides, sizeable expatriate or exile communities of Persian speakers live in adjacent countries, such as the Arab Emirates, Iraq, Turkey, Pakistan, and in other countries such as Russia, Australia, Japan, Israel, USA, Canada, and in various countries of western Europe. Most Iranians, Afghans and Tajiks have migrated there since the second half of the 20th century.

Reliable statistics for native speakers of Persian are not available for any of the three Iranophone countries. Conservative estimates would be around 50% for Iran, 30–35% for Afghanistan, and 80% for Tajikistan. Based on recent, partly estimated population figures (Iran 76 million, Afghanistan 30 million, Tajikistan 5 million), and taking into account also approximately 5 million expatriates (1–1.5 million in Turkey, 1 million in Russia [mainly Tajiks], 0.5 million in Pakistan [mainly Afghan Hazara in Quetta], 0.5–0.6 million in the USA, 0.4 million in Western Europe, etc.), this would give approximately 57 million native speakers of Persian.

In the three countries, however, many non-Persians are also fluent, or indeed native in Persian (esp. in Iran, see §1.4 on the sociolinguistic situation). A more realistic estimate of native Persian-speakers would probably be in the region of 75–80 million, with another 15–20 million that are fluent or near-native speakers of one of the three official variants of Persian. The estimates of Persian speakers living outside Iran are included in these figures but have their difficulties: statistics usually count holders of Iranian/Afghan/Tajik passports, whose second- or third-generation descendants may no longer be proficient in Persian. On the other

¹ I would like to thank Ramin Shaghaghi, Pejman Firoozbakhsh and Maximilian Kinzler for the inspiring discussions I had with them, and for the many helpful remarks on a draft of this article.

hand, naturalized immigrants no longer show up in the statistics but may still be native speakers. Besides, the Iranian community living in Turkey consists largely of Iranian (Azeri) Turks, some of whom would not qualify as native speakers of Persian.

The official status of Persian has been laid down for the Islamic Republic of Iran in the Constitution of 1979 (§15: “Fārsī is the … official language of Iran”) and for Afghanistan in the Constitution of 2004 (§16: “Pashto and Darī shall be the official languages of the State”). While the status of Fārsī as the official and national language of Iran has not been seriously challenged in the recent history of the country, the status of Darī has been more disputed in modern Afghanistan. Pashto, the native language of the Afghan majority population (of 50–55 %), was given political preference over Darī since the 1930s, and was called the “national language” of Afghanistan in the Constitution of 1964 (§3). It lost this privilege after the 1978 communist coup d'état, and has since been only one of the two official languages alongside Darī. Darī has consequently consolidated its *de facto* dominant position in administration and inter-ethnic communication, but the Pash-to-Tajik ethnic and cultural antagonism continues to permeate Afghan politics until today (see Nawid 2012).

The Tajik Constitution of 1994 specified Tajik as the state language and Russian as the language of inter-ethnic communication, but in 2009 a law was passed in parliament that confirmed Tajik as the sole official language, thus restricting the use of Russian (Landau and Kellner-Heinkele 2012: 177–179). The politics of linguistic Tajikization had been pursued systematically after independence in 1991, similar to “Uzbekization” in neighbouring Uzbekistan, etc. (Landau and Kellner-Heinkele 2012: 179–182). The Tajiks’ attitude towards Russian seems to be pragmatic nowadays, in that it remains an important language of education and academia, and is likewise important for those Tajiks who go to Russia as guest workers. The real challenge is for the Uzbek minority of about 15 % that is increasingly being marginalized, and perceived as the domestic agents of Uzbekization in Tajikistan (Landau and Kellner-Heinkele 2012: 184).

While the unity of Persian as *one* language in three Iranophone countries can be asserted on linguistic grounds and would be upheld by most speakers in the three countries (especially in Iran and Afghanistan), a tendency to develop a distinct linguistic identity may be postulated for Tajik. This is mainly based on the use of the Cyrillic script (as against the Arabic script used for Fārsī and Darī), on the strong lexical influence of Russian during the 20th century, and on the long-standing influence of Uzbek on the grammar of Tajik (especially on its northern dialects). While a counter-tendency to bring Tajik again closer to Fārsī has also been at work after the demise of the Soviet Union in the 1990s (see Landau and Kellner-Heinkele 2012: 182–184), it remains uncertain where Tajik-speakers will look for their linguistic affiliation and identity in the future.

1.2. Genetic affiliation

Persian is commonly grouped among the southern branch of the West Iranian languages, together with adjacent languages spoken in southwestern Iran like Lorī, Lārī, Kumzārī etc. The linguistic features that distinguish the Southwest Iranian languages from the Northwestern branch can be traced back to Middle Persian (3rd–7th century C.E.) and Old Persian (6th–4th century B.C.E.), e.g., the development of IE. *ǵ to Persian *d* in OP *dan-*, MP/NP *dān-* ‘know’, as against the northwestern (or better: non-southwestern) development to *z* in Avestan *zan-*, Parthian *zān-*, Kurdish *zān-*. Both Old and Middle Persian had their home in the historical province of Persis (Fārs) in southern Iran, from where the Achaemenid and Sasanian dynasties originated.

The SW-/NW-distinction was introduced by Tedesco in 1921, on the basis of features that distinguish Middle Persian as a typical SW, as against Parthian as a NW-Iranian language (spoken in the 3rd century C.E. in northern Mesopotamia); according to these features, Tedesco grouped also the modern W-Iranian languages as either “NW” (Tālešī, Semnānī etc.) or “SW” (Persian, etc.). The NW-/SW-distinction has been modified since the 1960s, and more recently been replaced by scalar models that group the W-Iranian languages on a “scale of Northwestern-ness”, in which Kurdish takes an intermediate position (see Paul 1998; Gippert 2007/08), but the SW-/NW-distinction, as two poles on a scale, can be maintained in this modified sense. The wanderings of the Iranian tribes throughout history, and the mutual influences their languages and dialects exerted on each other, were so intensive and complex that it is unlikely that any simple or reductionist model can do justice to the historical development of the W-Iranian languages.

From the outset (i.e., from Old Persian), Persian underwent stronger phonetic and morphological changes than the other, more conservative Iranian languages. For example, Middle Persian (between the 3rd and 7th centuries C.E.) had already lost almost all morphological distinctions of case and gender that had characterized Old Persian as an ancient Indo-European language, while these distinctions have still been preserved (albeit in reduced form) in many modern NW-Iranian languages until today. Persian has also undergone strong lexical influence from NW-Iranian, probably from Parthian or “Parthoid” dialects in which Iranian epics were transmitted.

1.3. History of the speech community

The evolution of Persian (or more precisely of “New Persian”) after the Arabic-Islamic conquest of Sasanian Iran (636–651 C.E.), and of its speech community, is long and complex. It is commonly held that the conquest dealt a fatal blow to Sasanian culture and its Middle Persian language. Subsequently, Arabic apparently

took over all official functions in the Caliphate and Persian only started regaining its old strength and literary, administrative etc. functions more than 200 years later, from the late 9th century onwards, in the course of a cultural reawakening under the Samanid dynasty that ruled NE-Iran.

A closer look, however, shows a different, and less simple, picture. Although Middle Persian ceased to be used as the official language of an empire around 650 C.E., it continued to be used on coins and in local administration well into the 8th century; for the 8th century MP documents from Tabaristan, see recently, e. g., Gyselen (2012). In the countryside of the vast area, many local strongholds of Zoroastrianism remained more or less intact for some time, where Middle Persian continued to be used as a Zoroastrian “church language”. The 8th/9th centuries saw a period of thriving Zoroastrian Middle Persian literature. This is likely to have been due also to the interaction of parts of the Zoroastrian clergy with other religions in the relatively “open-minded” atmosphere of the Caliphate during this period, when Mu’tazilism had a strong influence. The rise of Sunni orthodoxy in the late 9th century may have ended this.

A Jewish Iranian religious minority had lived in Iran since Achaemenid times, and wrote Persian already before the Muslim Persians did. The oldest Judaeo-Persian texts, two letters written in Hebrew script, were found in Dandan Uiliq (W-China) and are assumed to be from the 8th century C.E., which would make them the oldest extant New Persian texts. There is a flow of further Judaeo-Persian texts and documents from the late 10th century onwards. Since most of these texts are from SW-Iran, close to the still thriving centres of Zoroastrianism, they share certain dialectal features with Zoroastrian Middle Persian of S(W)-Iran (e. g., the use of the *ezafe* to introduce relative clauses), as against the new form of (Muslim) New Persian that was emerging in NE Iran at about the same time. The Zoroastrians and Jews referred to their languages with two historical versions of the same glottonym, MP *Pārsīg* / JP *Pārsī*, while the new Muslim form of New Persian in the NE was originally called *Darī* by its speakers (later the two glottonyms merged into *Pārsī-ye Darī*, or *Pārsī*). The use of these and other glottonyms (e. g., *Pahlavī*, which was the name of dwindling Parthian at the time) in early Arabic sources contributes to our understanding of the linguistic situation of the time, and the development of Persian (see the pioneering article by Lazard 1971, and a more recent account in Paul 2013b).

1.3.1. First steps in Persian literature

During the 10th century C.E., “Dari” New Persian written in Arabic script was gradually extended in its use in the Samanid Empire that ruled over NE-Iran, from lyric court poetry into further literary and scientific domains. Many of the first prose works written in Persian were translations from Arabic, e. g., the translation of the *Tārīx-e Tabarī* (“History of Tabarī”), or of the *Tafsīr-e Tabarī* (“[Qur’ān] com-

mentary of Ṭabarī”, both translated probably in the 960s C.E.); for an overview of the earliest NP prose texts, see Lazard (1963: 36–116). Although the genres, forms and contents of NP Muslim literature were novel to Iranian culture, elements of continuity from pre-Islamic Persian literature can be found in them.

The Iranian national epic *Šāhnāme* was composed by the poet Ferdousī (ca. 940–1019) probably on the basis of the “Abūmanṣūrī” prose version, which was derived from the (lost) Middle Persian *Xwadāy-nāmag*. The genre of New Persian wisdom/*andarz* literature, e. g. works like the *Siyar al-Molūk* (*Siyāsat-nāme*), written (in 1092) by the Seljuq vizier Neẓām ol-Molk, goes also back to pre-Islamic times, but the transmission from Middle to New Persian went partly through Arabic, or was at least substantially influenced by the way Arab writers like al-Jāhiẓ developed the *adab* genre in Arabic. Taken together with the preponderance of translations from Arabic among the earliest New Persian prose works, this shows how much NP literature owes to the classical Arabic tradition, how thoroughly it was influenced by it, yet how it managed to transport its own distinct, Iranian character into the new Islamic period. NP lyric poetry adopted the Arabic metric system, adapting it to the Persian syllable structure.

In the late 10th century C.E., Persian written in Arabic script was firmly established as a language of literature, science and administration at the Samanid court (NE Iran), and in the following (11th) century, Persian consolidated this status as the main language of the Ghaznavid and Seljuq Empires. Through the intensive mutual influence with Arabic language and culture since the 9th century, Persian had acquired the character of a deeply-rooted Islamic language, capable of expressing the most sophisticated contents of Islamic Civilisation. In science, especially in theology and philosophy, Arabic continued to be dominant for some time, but through epic works like the *Šāhnāme*, and the lyric of poets like Rūdakī (10th century), Farrokhī Sīstānī or Manūčehrī Dāmyānī (both 11th century), Persian had won a literary prestige that would pave its way as the dominant language, the upholder and disseminator of Islamic Civilisation, of the Eastern Caliphate. Through the military expansion of dynasties like the Ghaznavids, Seljuqs, later Ghurids and many more to come, Persian saw a massive extension of its geographical spread, into large areas of Central Asia, India, and Asia Minor, in the following centuries.

1.3.2. Persian as a lingua franca

Through the 11th–19th centuries, Persian was used as a language of literature, administration and diplomacy in these areas and beyond, as the lingua franca of the “Eastern Caliphate”, an area for which Fragner (1999) coined the term “Persophonie”, as an equivalent for the Persian term *Qalamrow-e zabān-e Fārsī* (i. e., “realm of the Persian language”). The linguistic dominance of Persian ended only in the modern period, when the colonial powers of Great Britain and Tsarist Russia

advanced into these territories, and vernacular (especially Indian and Turkic) languages rose to literary and national languages, though not without being deeply influenced by Persian.

The manner, intensity, and functions in which Persian was used over such a long period, varied greatly from region to region and from time to time. It was most prominent in India, where Persian had already gained a foothold as a court and literary language under the Delhi Sultanate (1206–1555). It saw its heyday under the Mughal dynasty (1526–1858), whose ruler Akbar the Great made Persian the official language of the Empire in 1582. Through the 17th–19th centuries, India is likely to have hosted far more (literate) readers of Persian than Iran itself (Cole 2002: 18). Persian was extensively studied and used also by Hindus who aspired to a career in state administration, and who produced excellent philological works on Persian, e. g. dictionaries (Alam 1998: 328–330). India was at that time actually not only hosting a Persophone diaspora, it was a centre of Persophony, on equal terms with Iran proper. Even under British rule, India retained its exceptional role for publications in Persian well into the 20th century; e. g., out of twenty-four known *Šāhnāme* lithographs from between 1846 and 1900, twenty-one are from India, and only three from Iran (Marzolph 2011: 70–71).

In the Ottoman Empire, Persian was highly esteemed as a language of learning, culture and poetry, and was part of the didactical “canon” of languages to be taught to Islamic students in madrasas, besides Arabic and Ottoman Turkic, up to the early 20th century (Kreiser 1993). That Persian was used as a language of trade and inter-ethnic communication also in regions north of the Ottoman Empire, on the Black Sea coast east of the Crimea, is shown by the Codex Cumanicus, a manuscript from about the 13th century that was probably composed by Christian missionaries and contains a Low Latin-Persian-Cuman Turkish dictionary (MacKenzie 1992).

In China, Persian never had the same prestige, or number of speakers or users, as it did in India or the Ottoman Empire, but in the administration of the (Mongol) Yuan Empire (1271–1368), it was the third official language (after Mongolian and Chinese), and remained an important language of study for the Chinese Muslims well into the Ming period (1368–1644) (Liu 2010: 89). The Mongols considered it as *the* language of the peoples living beyond China’s western borders, it thus became an important language of international diplomacy, see, e. g., already the letter of the Mongol Khan Göyük to Pope Innocence IV in 1246, written in Latin and Persian (Spuler 1985: 377). It was in 17th century China that the first Persian grammar was written in Persian (Kauz 2009).

The use of Persian over such vast areas and long periods was not uniform; although it declined in certain areas where vernacular languages rose to languages of administration and literature (e. g., in the Ottoman Empire after the 16th, or in India after the 18th century), it kept its high prestige as a language of classical Islamic culture and literature, in all areas, up to the early 20th century and beyond.

In the late 19th century, the Albanian poet Naim Frashëri (1846–1900) still wrote poetry in Persian, as did the Indian Muslim poet and philosopher Muhammad Iqbal (1873–1938), who by so doing aimed at a greater international dissemination of his ideas in the Muslim World than would have been possible through the medium of Hindustani/Urdu (Rypka 1968: 371).

Although the expansion and development of Persian as a lingua franca was always closely aligned to developments in Iran proper, it must be noted that it was generally independent of any political, financial, or other institutional support from Iran's ruling dynasties. A concrete measure like the introduction of Persian as a language of administration in the Mughal Empire (in 1582) was, of course, a deliberate political decision, but it was not an ideological measure to promote a language in the modern sense. Rather, it was a practical decision: Persian was considered the language available and most suitable for meeting the requirements of administering a large empire, and its high prestige as a language of court, administration, literature, learning and Islamic culture further recommended it for this role.

Interestingly, almost none of the many dynasties that used Persian as an official language throughout the 11th–19th centuries, including the dynasties ruling Iran proper, were of Iranian origin. Most (like the Mughals) were of Turkic extraction, and assimilated to the Persian that they regarded as more suitable for administration, or more prestigious, than their own vernacular language. Paradoxically, while the rulers of dynasties outside Iran like the Mughals, often spoke Persian, the Safavid rulers of Iran proper often talked Turkish at court. The Shia-Sunna antagonism that emerged in the region through the rise to power of the Safavid dynasty in 1500 and the following Shiization of Iran, had no effect on the cultural and literary exchange between Iran and the Persianate world: Persian was already too well-established to be affected adversely by this.

1.3.3. The modern period

The demise of the “Persian cultural Commonwealth” of the Eastern Islamic World was brought about not only by regional vernaculars like Hindustani that became established as literary languages from the 18th century onwards, but mainly by the two colonial powers of Russia and Great Britain. Both empires extended their territories considerably during the 19th century, at the cost of Persophone regions of Central Asia and India, and introduced new forms of government and administration and new means of communication there, replacing Persian as a language of administration and communication by Russian, English, and Hindustani. Persian was abolished as the official language of administration in India by the British rulers in 1835 (Bonakdarian 2006: 26).

Iran proper did not remain unaffected by the advances of the two colonial empires. In the early 19th century, Iran lost most of its Caucasian territories to

Russia, and in 1857, it finally had to waive its claims to the province of Herat that became integrated into the new Afghan “buffer” state between Russia and British India. Some members of the Iranian ruling elite realized that in order to survive, Iran was in dire need of reform and modernisation. In the cultural sphere, certain Iranian literati started criticizing the bombastic and verbose “Indian” style of poetry that had held sway at the Safawid and Mughal courts through the 16th-18th centuries, and advocated the simpler “Khorāsāni” and “Iraqi” styles of classical authors. This movement was called *bāzgašt-e adabī* “literary return”. Its most important protagonist, in poetry and prose, was the writer, poet, vizier, and reformer Qā’em Maqām Farāhānī (1779–1835; see Rypka 1968: 335–336). One of the best-known samples of the new, simpler Qajar prose works composed in the late 19th century are Šāh Nāṣeroddīn’s diaries on his three journeys to Europe between 1873 and 1889.

In the 20th century, Iranian prose writers started including passages of colloquial Persian in written form for the first time, e.g. Mohammad ‘Alī Čamālzāde in his collection of short stories *Yekī būd, yekī nabūd* (1921). With the adoption of European genres like short story or romantic novel, “realistic” writing became the dominant paradigm in Iranian prose literature. The rapid changes in style and grammar of 20th century written Persian, however, were due rather to a sudden breakup of literary conservatisms, than to an actual change of the language structure; a closer look at the grammar of Persian will show that many innovations and colloquialisms could be found in written Persian already well before the 19th century, and that the structure of Persian changed steadily over time rather than rapidly in response to new genres.

There is no universally accepted periodization of the history of Persian. National Iranian philology prefers a subdivision into styles (*sabk*) that are identified with certain regions where these styles developed, or flourished most prominently. Although one can identify “classical” periods of these styles (Khorasani 9th–11th centuries, Iraqi 12th/13th centuries, Indian 16th–18th centuries, followed by the Return [*bāzgašt*] movement), they are in principle independent of time, and any poet could, and can, write poetry in any of these styles at any time (on this and alternative subdivisions, Rypka 1968: 112–119). Western linguists have proposed various subdivisions. The one advocated here follows Lazard (1963: 24, shared also by Khanlari), into Early (8th–12th centuries), Standard (13th–19th centuries), and Modern High New Persian (19th century—). It should be noted that “classical” is a literary term covering classical authors of various times (Ferdousī, Sa‘dī, Hāfez, Jāmī etc.), and is, therefore, of limited linguistic value, since the language of these authors is by no means uniform (see Paul 2002). In terms like: “Early New Persian and Classical Persian”, the latter may, however, be used to denote the earliest stage of Standard Persian, i.e. the language of the classical authors of the 13th–15th centuries.

1.4. Sociolinguistic situation

As the official or state language of Iran, Afghanistan and Tajikistan, Persian enjoys a privileged position in the state institutions of these countries, and in Iran, it receives further support from a national language academy (*Farhangestān*). In all three countries, sizeable non-Persian ethnic minorities (in Afghanistan, a majority) have been resident for a long time, among whom bilingualism with Persian as a second language is more or less wide-spread. Persian naturally advances among the non-Persian minorities through administration, the educational system, and the mass media. At the same time, in all three countries, Persian is under pressure from other languages, in Afghanistan (expectedly) from Pashto, in Iran from Azeri Turkic, and in Tajikistan from Uzbek.

1.4.1. Bilingualism

The ethnic and linguistic communities' command of their own language(s), dialects, and of Persian, and thus the grades of bilingualism, vary greatly from case to case. While the predominantly Sunni minorities of Iran, the Kurds and Baloch, adhere to their ethnic languages quite firmly, reinforced by their distinct religious identity, the large Shia linguistic minorities of northern Iran (Gilan and Mazandaran), who number approximately 3–4 million, have already been largely assimilated to Persian, rendering the future demise of the languages very likely (Shahidi 2008: 295). The same is true of many other smaller language or dialect communities of NE and Central Iran like Semnānī, Vafṣī, etc. Among Shiī Kurds of SW Iran, e. g. in the city of Kermanshah, Kurdish is still strong but has been more mixed with Persian, and with local idioms like Lakī or Kalhorī, than the “Sunni” Kurdish of Sanandaj.

Idioms like Gīlakī, Māzandarānī or Semnānī that are clearly languages in their own right linguistically, are not granted the status of a “language” in the Iranian world-view. They do not possess a written literature, or literary tradition, comparable to that of Persian, are mainly used orally and in informal situations, and are, therefore, called “dialects” (*gūyeš*). This does not do justice to Māzandarānī, which can boast of a rich and relatively old written literature, going back at least to the early 19th century, and possibly as far back as the 11th century, if Māzandarānī is to be identified with “New Ṭabarī” (Borjian 2009). But the knowledge about this remains restricted to specialists and folklorists. It also attests to the unique status and extraordinary prestige of Persian, which is perceived as *the* national and literary language *par excellence*, and in comparison to which no other language is even granted the name of “language”. The Kurds and Baloch, whose distinct (Sunni) ethnic-religious identities are accepted by Shiī Iranians, are, however, exempted from this practice and their idioms are usually referred to as “languages”.

The case of Azeri Turkic is exceptional and contradictory. Although the Azeri Turks of Iran are Shiites and firmly grounded in the Iranian-Shii cultural and national identity, and although they still seem to take for granted the “superiority” of Persian over Turkic as a language of literature and culture, the position of Azeri Turkic is very strong in the official Turkophone provinces (W- and E-Azerbaijan, Ardabil), and Persian is less well-known and used there than in all other Shii regions of Iran. Azeri Turkic even seems to be advancing geographically and demographically in various provinces like Zanjan, Qazvin, Hamadan or Gilan, at the cost of minority languages like Tālešī (most Tāleš, especially the majority who are Shii, seem now to be Turkicized), Tātī and Gilakī, but also of Persian. For the large Azeri community of Tehran, a study says that although there is an ongoing shift from Azeri-Persian bilingualism towards Persian, the vitality of the Azeri language is not in danger (Bani-Shoraka 2005: 197–202). Dialects of Azeri Turkic are the undisputed “L1” in all Turkophone provinces and beyond, while in Gilan and Mazandaran, but also in Kermanshah and other regions of Iran, part of the local (ethnically non-Persian) population would probably call Persian their first mother tongue.

From the “horizontal” (or geographical) bilingualism, a “vertical” one may be distinguished, indicating the coexistence of various historical forms and styles like classical literary Persian (of various [sub-]forms), modern literary high Persian, and modern colloquial Persian. Each native speaker can, depending upon his education, switch more or less freely between these styles, according to situation, addressee, place, time, and other factors; mixed or intermediary forms, like high Persian with elements of colloquial influence, are also possible (see Paul, forthcoming). The mixture of styles and registers, especially the continuous normative influence of Classical Persian on the modern written language, makes it difficult to define a fixed or homogeneous standard of contemporary Persian. SNP may rather be defined as a flexible system of rules that fall more or less within the limits of a standard.

While the social status of many Iranians may be guessed from their accent, or from the way and extent they use colloquial style, another, totally different system of rules determines and permeates Iranian social interaction, and thus also linguistic behaviour, namely the system of politeness called *ta ‘ārof*. It determines the way that any Iranian talks to or about a person of equal or superior social position, requiring the use of certain forms, words or expressions which are different from those he would use if talking to a child, or a person of (visibly) inferior social position. *Ta ‘ārof* is variously described as—neutrally—the “system of politeness specific to Iran”, or—negatively—as the “most hypocritical exaggeration and distortion of a system of politeness”. It affects mostly the lexicon (e. g., using verbal constructions like *tašrif āvardan* “to come”, literally “to bring honouring”, or the noun *bande* “slave”, for the 1SG pronoun “I”), but also the grammar (3PL, instead of 3SG, for talking about a “superior” person, e. g. *īšān hastand* “he is present”, literally “he/they are present” (see Beeman 1986: 151–162 for an overview).

1.4.2. Education, administration

In the educational system of Iran, Persian is the only language of instruction, but local minority languages seem to be tolerated in schools to a certain extent. In administration, Persian is again the only official language, but other languages may be used pragmatically, e.g., if a person at court does not speak Persian well enough. The printing of books and journals in minority languages is not discouraged or prohibited as such, on linguistic grounds, and there are no limitations, either, for the use of minority languages in public. Under the pressure of Persian as a language of education and the media, an unforced assimilation of many minority languages can be witnessed, often with the (parents') motivation not to hamper the children's prospects for their later professional career through a "provincial" accent.

1.4.3. National academies

Of the three Iranophone countries, only Iran has a fully-fledged national language academy (*Farhangestān*), which is engaged in various philological (editorial, lexicographical etc.) activities, and in an ambitious terminology programme, coining thousands of Persian terms in over sixty fields of knowledge every year, with the aim of replacing scientific terms of foreign (English, Arabic) origin. For an overview of the *Farhangestān*'s history since 1935 and its main activities, see Jazayery (2000).

In Afghanistan, a Pashto Language Academy (*Paštō Tōləna*) was founded in 1974, with the aim of promoting the "under-developed" national language. Since Persian had always been the dominant language of administration and culture in recent Afghan history, obviously no need had been felt to support Persian through an institution of its own until today. In Tajikistan, there is a national Academy of Science going back to the Soviet period, whose philological section is also active in linguistic matters from time to time, but there is no institution comparable to the Iranian *Farhangestān*. In its cooperation with Afghan and Tajik scholars, the Iranian *Farhangestān* clearly tries to cover commitments that go beyond Iran and Fārsī alone, and to take responsibility for "Persian" as a cover term for Fārsī, Darī, and Tojiki.

2. Remarks on grammars and reference works

No comprehensive grammatical description of the Persian of Iran (Fārsī) exists that takes its historical development into consideration, and complies with modern linguistic standards. The historical grammar of Khānlarī 1986, parts of which have been translated into English (1979), remains within the framework of tradi-

tional philology. Lazard 1957 (English translation 1992) is the standard reference grammar up to the present. It describes modern standard Persian, but takes into account also colloquial registers, and important rules of Classical literary Persian. In his 1989 sketch grammar, Lazard provides an updated, condensed version of his views, shifting the emphasis in certain details.

All other grammars of Persian lag behind Lazard in comprehensiveness, precision, and method. Phillott (1919) remains a rich collection of linguistic material, but lacks linguistic method. Jensen (1931) offers a remarkable coverage of syntax, uncommon for its time. Windfuhr (1979) is still a valuable research report, with a comprehensive bibliography. A reliable and detailed recent Persian reference work on Persian grammar is the dictionary, alphabetically arranged, by Țabătabāyī (2016).

The first comprehensive grammar of Tajik in a “Western” language is Perry 2005, published almost 50 years after Lazard’s very useful study of 1956. The first systematic comparative grammar of Fārsī and Tajik was given by Windfuhr and Perry (2009). Methodologically, this study includes many ad hoc examples from both idioms to illustrate the grammatical rules, rather than authentic examples from actual linguistic corpora. Besides the titles mentioned, only works on specific periods or variants of Persian, or on specific parts of its grammar, exist, three important ones of these being Lazard’s seminal study on Early New Persian (1963), Paul’s grammar of Early Judaeo-Persian (2013a), and Pisowicz’s comprehensive historical phonology of Persian (1985).

3. Phonology

3.1. The phonological system

The phonemic system of modern Standard New Persian (SNP) of Iran includes the following system of 26 consonants (including two semi-vowels) (Table 1) and 6 vowels (Table 2) (the symbols in the tables represent those of standard transcription, rather than the IPA).

Table 1: Consonant inventory

		labial	labio-dental	alveolar	post-alveolar	palatal	velar	uvular	glottal
plosive	voiceless	p		t			k	q	(‘)
	voiced	b		d			g		
fricative	voiceless			f	s	š		x	h
	voiced			v	z	ž		(γ)	
affricates	voiceless					č			
	voiced					j			
liquid	nasal	m		n				(ŋ)	
	lateral			l					
	trill			r					
semi-vowels	(glides)	(w)					y		

Table 2: Vowel inventory

	front	back
closed	ī	ū
middle	e o	
	a	
open	ā	

Diphthongs: *ey*, *ow*, *ay* (short; *ay* only in Arabic loanwords); *āy*, *ūy* (long). If *ow* is regarded as a biphoneme, a “limited” phoneme [w] must be postulated for SNP (Pisowicz 1985: 24–27).

3.2. Remarks

The phonemic status of some of these sounds, particularly those in brackets, is controversial. The glottal stop [‘], whose phonemicity is disregarded by some scholars, is considered phonemic here, following Pisowicz (1985: 47–51). The voiced fricative [γ] is understood as a positional allophone of uvular [q] (both sounds constitute one phoneme, but there is a discussion about which is its main representative, see Pisowicz 1985: 42–44). The glide *w* occurs only in limited phonetic contexts, like in the diphthong [ow] (Pisowicz 1985: 24–26). The nasal [n] has an allophone [ŋ] before the velar or uvular stops [g, k, q] (Pisowicz 1985: 18).

It should be noted that the above system represents modern standard Persian (SNP) in current Tehran “standard” pronunciation. Different styles, or regional varieties of Persian outside Tehran, may affect the above system in details. For

example in certain registers the glottal stop may be dropped in some phonetic contexts, without, however, giving up its phonemicity (see below). On the other hand, in careful and slow Tehrani pronunciation, and in certain regions of Iran (e. g., Fārs) generally, the [γ] retains its phonemic character, being pronounced differently from [q] in all positions.

3.3. Historical changes

The changes of the phonemic system in the history of Persian have been scarce in consonantism, but more substantial in vocalism. The most important consonant changes include:

- The [w] of Early and Classical Persian changed to modern [v], which did not exist in Classical Persian; In Tajik, [w] still exists today, besides [v].
- In Early and Classical Persian, the distinction of q/γ was still phonemic, as it is still today in many regions of Iran, and in all Afghanistan and Tajikistan.
- In Early New Persian, [β] and [δ] existed as allophones of the voiced stops [b] and [d] (Pisowicz 1985: 107–111).
- In Early and Classical Persian, [xʷ] was a (single) phoneme, which was de-labialized to [x] in later Persian (Pisowicz 1985: 121–123)

3.4. Vowel system

The Early New Persian (ENP) vowel system, where quantity was phonemic, developed into the modern system, in which the (historical) long vowels are still longer in articulation, but in which the main load of the distinction is on vowel quality. In Tajik, vowel quantity has been given up, and the development of the vowels differs considerably from the Fārsī one. The following scheme gives an overview of the main changes:

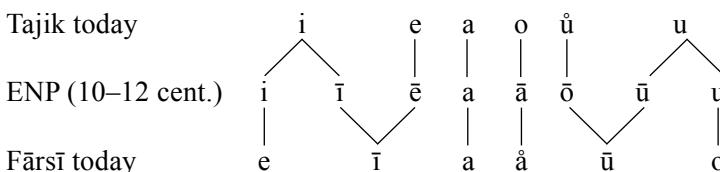


Figure 1: The historical development of the Early New Persian vowel system

This again describes only the two main vowel systems of Tehrani SNP and Literary Tajik. There are regional variants with divergent developments. In Afghanistan, and in most variants of Khorasan Fārsī, ENP [ē] is preserved as such (instead of SNP [ī]); in certain N-Tajik dialects, [ū/u] has developed into front closed rounded [ü] (instead of u), and [ō] into a stable, short [u] (Lazard 1956: 128).

3.5. Other phenomena

Besides the phonemic features described so far, there are many processes of co-articulation and allophonic variation in the pronunciation of modern colloquial Persian, that do not (or only marginally) touch upon the phonemic system of SNP. A selection of these includes:

- a) SNP *ā* before nasal (*m*, *n*) > CollP *ū* (e. g., *-ān* > *-ūn*, Šādeqī 1984).
- b) The short vowels [a, e, o] are unstable and may undergo changes in unstressed position, e. g. *kelīd* ‘key’ > *kilīd*.
- c) The glottal stop ['] and the ‘ayin in Arabic loanwords, if followed by a consonant, are realized as a lengthening of a preceding vowel, e. g. *ma'lūm* > *maalūm* ‘known’.
- d) [q/γ] is fricativized before a voiceless fricative, e. g. *raqṣīdan* [raksīdan] ‘to dance’, exceptionally also before [t], e. g. in *waqt* [vaxt] ‘time’ or *eytešāš* [extešāš] ‘riot’.
- e) [k] and [g] are palatalized before the vowels [a, e, i] e. g. *kār kardam* ‘I worked’.

3.6. Influence of Arabic

The massive lexical impact of Arabic on Persian during the first centuries of Islam notwithstanding, the influence of the Arabic sound system on the Persian one has remained limited, the only phonemes clearly borrowed into Persian being [q] and ['] (on the borrowing of [q], see Šādeqī 2006–07). Although pronounciations of Arabic emphatic and pharyngeal sounds like [d̪] or [h] may have existed in “learned” registers of Persian at certain times, they did not exert a lasting influence on the Persian sound system.

3.7. Geminates

In New Persian, geminate consonants are phonemic, but most distinctions are imported from Arabic through the many Arabic loanwords with geminate consonants, e. g. NP *māde* ‘female’: *mādde* (<Ar.) ‘matter; material’; or with a geminate in the Persian word, see NP *kore* (<Ar.) ‘globe’: *korre* ‘foal’.

3.8. Syllable structure, consonant clusters

The following six types of syllables occur in modern Standard Persian: (C)V, (C)VC, (C)VCC, with V being a short or long vowel. If one considers (as some scholars do) the syllable types with initial vowel as starting with ['], the number of types would decrease to three (CV, CVC, CVCC). Following Pisowicz (1985: 50–51),

in initial position, the glottal stop may be understood not as a distinctive, but only as a delimitative feature.

NP syllable structure does not allow initial consonant clusters. Zoroastrian Middle Persian had initial clusters (ZMP *stārag* ‘star’, *fradāg* ‘tomorrow’), but in Manichaean MP, a prosthetic [i-] preceded clusters with *s-/š-* (MMP *istārag* ‘star’; but ‘tomorrow’, MMP *fradāg*). Early New Persian exhibits variation between prothesis and epenthesis to break the cluster (e. g., ENP *sitāra/istāra* ‘star’, Lazard 1963: §105; EJP *istāra* ‘star’, Paul 2013: §47.d). Clusters with consonants other than *s-/š-* mostly took an epenthetic vowel (EJP *firdā* ‘tomorrow’, Paul 2013a: §40). In later Persian, epenthesis prevailed in all clusters (NP *setāre/fardā*). In modern NP, loanwords with *s-/š-*-clusters again receive a prosthetic vowel (e. g., *estekān* ‘glass’ < Russ. *stakan* ‘glass’, *eslāyd* ‘slide’), while those without *s-/š-* show epenthesis (*kelās* ‘class’, *terāmvāy* ‘tramway’); see also Lenepveu-Hotz 2011.

3.9. Hiatus

If a word ending in vowel takes an ending or clitic that also starts with a vowel, a hiatus is normally avoided through a glide, mostly -*y-*, e. g. *xāne-y-e man* “my house” (-*e* ezafe), *pā-y-am* “my foot” (-*am* possessive suffix), more seldom -*v-*, e. g. *karre v-o nān* “butter and bread” (*o* “and”). The clitic copula, however, allows either hiatus or -*y-:* *īnjā-(y-)am* “I am here” (the form with -*y-* being rather colloquial).

A hiatus after -*a* or -*ā* may alternatively be avoided by contraction, e. g. *pā-m* “my foot”, *īnjā-m* “I am here”. After other vowels, only the possessive suffix allows contraction, hence the distinction between *pahlū-m* “beside me” vs. *pahlū-y-am*, meaning either “beside me” or “I am beside”.

3.10. Stress

Stress generally falls on the final syllable of each stem, e. g. *ketāb* ‘book’, *ketābi* ‘bookish’, but *rāft-am* ‘I went’. In the last word, -*am* is a verbal suffix that remains unstressed. The extent to which grammatical suffixes and clitics attract stress varies considerably (see Kahnemuyipour 2003 for a comprehensive theoretical analysis). In general, clearly inflectional items do not bear stress, for example person and number agreement suffixes on verbs, or clitics such as the indefiniteness suffix -*ī*, *ezafe*, clitic pronouns, postposition =*rā*, the short copula, etc. Derivational affixes, as -*ī* in *ketābi* ‘bookish’, on the other hand, are stress-bearing items. The nominal plural suffixes -*ān* and -*hā*, the comparative suffix -*tar* and the superlative suffix -*tarīn* are also stress-bearing items (e. g., *ketābhā* ‘books’, *bozorgtār* ‘bigger’). The verbal prefixes *mī-* (present indicative, durative), *be-* and *na-/ne-* (negation) are also always stressed, e. g. *mī-raft-am* ‘I used to go’, *nā-raft-am* ‘I didn’t go’.

(Pisowicz 1985: 55–56). If negation and *mī-* occur together, the former takes the stress: *né-mī-raft-am* ‘I did not go’.

Stress can be distinctive, see *ketābī* ‘bookish’: *ketāb-ī* ‘a book’ (with the indefiniteness suffix *-ī*, §4.5.); see also the distinction between simple past and present perfect (§6.8.).

4. Nominal morphology

4.1. Nouns

New Persian nouns distinguish number and definiteness, but there is no morphological case system, nor a distinction of animacy or gender; in Arabic loanwords, however, the Arabic gender distinction (M/F) has been imported into New Persian to some extent. The NP bare noun is generally neutral with respect to definiteness and number:

ketāb ‘book, a book, the book, books’

4.2. Distinction between nouns and adjectives

The distinction between nouns and adjectives is not always clear-cut in NP, e. g. many adjectives can be substantivized (*bozorg* ‘big’ > *bozorg* ‘the big (one)’, pl. *bozorgān* ‘the big (ones), the elders’); certain nouns, in turn, may be adjektivized *ad hoc* through an adverb of measure (*xeilī* ‘very’), or the comparative suffix *-tar* (see the noun *mard* ‘man’ in: *xeilī mard ast* ‘he is very man(ly)’; *mard-tar ast* ‘he is more manly, more (of a) man’).

4.3. Number

A NP simple noun may be understood generically, and be used instead of an English plural:

- (1) *īnjā sīb ziyād ast*
here apple.SG much COP.3SG
'Here there are many apples'
- (2) *man ketāb mī-xaram*
I book.SG DUR-buy.PRS.1SG
'I buy books'
- (3) *az peste xoš-am mī-y-āyad*
from pistachio.SG good-my DUR-GLIDE-come.PRS.3SG
'I like pistachios'

The simple form of the noun is also used for plural nominal predicates:

- (4) *ānhā kārgar-and*
 they worker.SG-COP.3PL
 ‘They are workers’

In most other cases, the plural is marked by one of the stressed suffixes *-ān* or *-hā*, by an imported Arabic suffix (*-āt*, *-īn*, etc., §4.4.), or by an Arabic ‘broken’ plural form. The distribution and functions of these distinct plural formatives may be summarized in the following way.

The stressed plural suffix *-ān* goes back to a MP plural oblique ending *-ān*. In Classical and modern Standard Persian, it is used for animate human nouns (e. g., *zanān* ‘women’), and for a small number of other nouns (e. g., *deraxtān* ‘trees’, *dastān* ‘hands’). Stressed *-hā* goes back to the MP adverbial ending *-ihā* (the EJP plural suffix is still *-ihā* in most texts) and denotes the plural of inanimate, and of animate non-human nouns (e. g., *saghā* ‘dogs’).

In modern colloquial Persian, *-hā* has almost completely replaced *-ān*, also for human nouns, e. g. *mardhā* ‘men’ (*-hā* with human nouns appears in prose literature already in the 16th century, see Paul 2002: 31). Certain nouns have developed different meanings in SNP, depending on the use of *-ān* and *-hā*, e. g. *sarān* ‘heads, leaders’ vs. *sarhā* ‘heads (i. e., the body part)’. In some cases, the use of *-ān* and *-hā* is idiomatic, e. g. *xānomhā va āqāyān* ‘ladies and gentlemen’ (**āqāhā* is not possible).

Unlike English, Persian plural forms can be used with mass nouns and infinitives, to denote large quantities or repeated actions:

- (5) *īn pūl-hā*
 DEM money-PL
 ‘these lots of money’
- (6) *īn raftan-hā*
 DEM going-PL
 ‘this going and going’

In certain contexts, *-hā* may make a noun phrase definite:

- (7) *čahār (tā) doxtar*
 four NUM daughter (for numerals, see §4.6.)
 ‘four daughters’ (ordinal numbers occur regularly with the singular, §4.7.)
- (8) *čahār tā doxtar-hā*
 four NUM daughter-PL
 ‘the four daughters’

The SNP verb obligatorily agrees in number with a plural subject of a sentence only if this denotes an animate being. With inanimate subjects, the verb often remains in the singular:

- (9) *čerāy-hā xāmūš šod*
 lamp-PL extinguished become.PST.3SG
 ‘The lamps went out’ (Lazard 1992: §170, p. 179)

4.4. Arabic plural forms in NP

Together with the many Arabic loanwords, the following Arabic plural forms were also introduced to Persian:

- i. *-īn* (< Ar. GEN.-ACC.PL.M): in SNP, after a limited number of Ar. nouns denoting mostly masculine humans, e. g. *mo‘allemin* ‘(male) teachers’, but also humans of both sexes, e. g. *mosāferīn* ‘travellers’.
- ii. *-ūn* (< Ar. NOM.PL.M): after a very small number of Ar. nouns ending in *-ī*, e. g. *rūhāniyūn* ‘clergymen’ (neither *-īn* nor *-ūn* have preserved the original Ar. case function).
- iii. *-āt* (< Ar. F.PL): regularly after Ar. (F) nouns in *-e*, e. g. *kalamāt* ‘words’, *moqaddamāt* ‘preliminaries’ (sg. *kalame*, *moqaddame*), but also after certain other Ar. nouns, e. g. *entexābāt* ‘elections’, *touzīhāt* ‘explanations’ (sg. *entexāb*, *touzīh*).
- iv. ‘broken’ Ar. plural forms, e. g. *afkār* ‘thoughts’ (sg. *fekr*).

The use of the plural ending *-āt*, and of the broken plurals, has been extended from Arabic loanwords to originally Persian words, and become productive to some extent in SNP, e. g. in: *deh-āt* ‘villages’, *kārxāne-jāt* ‘factories’, *sabzī-jāt* ‘vegetables’; the additional *-j-* goes back to a MP *g* that developed into *j* in MP loanwords in Arabic. The broken plural is found in (originally Persian) ethnonyms like *Akrād* ‘Kurds’ (sg. *Kord*), or in *darāvīš* ‘dervishes’ (sg. *darviš*). Arabic plurals sometimes receive an additional, “pleonastic” Persian or Arabic pl. ending, e. g.: *tazāhorāt-hā* “demonstrations”, *javāher-āt* “jewels” (*tazāhor-āt* PL, *javāher* broken PL)

In modern colloquial Persian, all Arabic plural forms have (like Persian *-ān*) been largely replaced by *-hā*, e. g. *mosāferhā* ‘travellers’, *touzīhāt* ‘explanations’, *kārxānehā* ‘factories’. In those registers where both plural options (Arabic and Persian) are possible, this may cause a semantic distinction, with the Arabic plural form often taking a more collective, abstract, or specialized meaning, e. g., *fekrhā* ‘thoughts’, but *afkār-e omūmī* ‘public opinion’; or *moqaddamehā* ‘(printed) forewords’ against *moqaddamāt* ‘(abstract) preliminaries, first steps’ (from sg. *fekr* ‘thought’, *moqaddame* ‘foreword’). For a comprehensive study of Arabic loanwords in Persian that end in *-a(t)*, see Perry (1991).

The mixture of Arabic nouns with Persian endings, and vice versa, shows that although most Arabic loanwords can still be easily distinguished from originally Persian words in the Persian lexicon today, they have been firmly integrated into the latter; together with them, the Arabic plural forms and rules have also found their place, with adjustments and simplifications, within Persian nominal morphology.

The Arabic Dual (GEN-ACC) ending *-eyn* has been borrowed into Persian together with a limited number of nouns denoting pairs, e. g. *vāledeyn* ‘parents’. This ending has not become productive in Persian.

4.5. (In)definiteness

An unstressed *-ī* (< CP/ENP *-ē*) may be suffixed to a NP noun, giving it the meaning of indefiniteness and singularity: *ketāb-ī* ‘a book, some book, any book; one book’. Although both aspects of this morpheme are given separate names by Persian grammar (*yā-ye vahdat* ‘ī of one-ness’ / *yā-ye nakere* ‘ī of indefiniteness’), they cannot really be separated in modern Persian.

The singular component of the meaning of this suffix is evidently overwritten when the *-ī* follows a plural noun to yield the meaning of indefinite quantity, e. g.: *zanhā-(y)ī* ‘some/certain women’. The *-ī* of indefiniteness/unit can be replaced by *yek* ‘one’ (*zan-ī* = *yek zan* ‘a woman, one woman’); in modern colloquial Persian, both occur often together: *yek mard-ī* ‘some man, a man; one man’. *Yek* may also occur together with the plural and *-ī*, e. g.: *yek mardhā-(y)ī* ‘some/certain men’.

A phonetically identical suffix *-ī* also indicates that a noun is indefinite, but it specifies it by “announcing” a following attribute that defines it (e. g., a relative clause, §8.2.; this *-ī* is called *yā-ye ešārat* ‘ī of reference’). It occurs regularly before relative clauses in modern Persian, but less often in the Classical language. It is probably of a different etymological origin from the indefinite *-ī*, with which it seems increasingly (but not yet completely) to merge, to form one morpheme (Lazard 1966, 1989: 275–276).

The marker of definite direct objects, *=rā*, and its occurrence together with indefinite *-ī*, is discussed in §5.5.

Definite *é*

Neither Classical nor Standard Persian have a definite article. In modern colloquial Persian, there is a stressed suffix *-é* that makes a noun definite, e. g.: *pesan-é* ‘the boy (in question)’. It occurs only with singular nouns, and together with demonstratives (e. g., *īn pesaré* ‘this boy’, Lazard 1992: 73–74). It may be followed by *=rā*. It does not occur together with indefinite *-ī*, but with specifying *-ī*, e. g. *ān pesaré-ī ke dīdam* ‘that boy whom I saw’.

4.6. Adjectives

SNP adjectives, if they qualify a noun, follow it with the ezafe (but see §7.1.3.) and are always in their basic (sg.) form, even when the noun is in the plural, e. g.:

- (10) *xānehā-ye bozorg*
 house.PL-EZ big.SG
 '(the) big houses'

A comparison is expressed with the comparative suffix *-tar* and the preposition *az* ‘from’:

- (11) *az man bozorg-tar ast*
 from me big-COMP COP.PRS.3SG
 '(he) is bigger than me'

4.7. Numerals and classifiers

In all variants of NP, cardinal numbers are used with the singular of nouns, e. g. *do mard* ‘two men’, *dah xāne* ‘10 houses’. In all variants of NP, there is a class of ‘numerative’ nouns (or ‘classifiers’) that may be inserted between a number and the noun that is counted, e. g. SNP *yek dast lebās* ‘one suit’ (lit., ‘one hand (of) suit’), *se nafar sarbāz* ‘three soldiers’ (‘three person soldier’). In CP and high registers of SNP, there were (or still are) dozens of these nouns whose usage depends on the semantic class of the counted noun, e. g. *dast* ‘hand’ for clothes, or *ra*’s ‘head’ (< Ar.) for cattle.

In modern spoken Fārsī and colloquial registers, the number of numerative nouns has been reduced to three (*nafar* ‘person’ for persons; *tā* ‘piece’ and *dāne* ‘grain’ for things). Their use is not obligatory, but they are often used. In colloquial registers, *tā* tends to be generalized also for persons (*dah tā mard* ‘10 men’), as does *dūne* (< SNP *dāne*), but the latter only for one-ness (*ye dūne mard* ‘a man’; see Lazard 1992: §75, Moinfar 1980, and Nešāt 1990).

4.8. Pronouns

The full paradigm of pronouns of Standard Persian are given in Table 3. Pronouns do not inflect for case.

Table 3: Standard Persian pronoun paradigm

Personal pronouns	
<i>man</i>	‘I’
<i>tō</i>	‘you’
<i>ū</i>	‘he, she, it’ (ENP/CP also: <i>vey</i>)
<i>mā</i>	‘we’
<i>šomā</i>	‘you’
<i>ānhā</i>	‘they’

When followed by the direct object marker *=rā*, *man* ‘I’ may lose the final nasal (*ma=rā* ‘me’); alternatively, in colloquial Persian, where *=rā* is shortened to *-o*, the *-n* is preserved (*man-o* ‘me’). The 2SG polite address is 2PL *šomā*. In modern polite language, a 3SG polite form *īšān* has developed, which goes back to the former ENP 3PL pronoun *ēšān* ‘they’ (that was replaced by *ānhā*, which was originally the pl. of the demonstrative pronoun *ān* ‘this’). Polite 3SG *īšān* takes a pl. verbal form (e.g., *īšān hastand* [he.PL COP.3PL] ‘he is here’). There is also a form of “halfway/reduced” politeness, by using *īšān* with a sg. verb, e.g. *īšān goft* ‘he said’. The noun *bande* ‘slave’ is used as a polite form for the 1SG pronoun (as it is considered polite to belittle oneself). This occurs with 1SG verbal forms (e.g. *bande raft-am* [slave go.PST-1SG] ‘I went’). In CP, *bande* in the sense of “I” was still used with a 3SG verb in the 13th century, while polite *ēšān* occurs at least since the 16th century (Paul 2002: 31–32).

In colloquial Persian, the pl. pronouns *mā* and *šomā* may take the pl. ending *-hā* (*māhā* ‘we’, *šomāhā* ‘you [PL]’), probably by analogy to 3PL *ānhā*, or in the case of *šomā*, to distinguish it from the polite form *šomā*, which is used for sg. address.

4.9. Pro-drop

New Persian is a pro-drop language, in which not only a subject may be dropped, e.g. *raftam* [go.PST.1SG] ‘I went’, but—in appropriate contexts—also a direct object pronoun:

- (12) *mašq=at=rā neveštī? bale, neveštam*
exercise=your=OBJ write.PST.2SG yes write.PST.1SG
‘Did you write your exercise? Yes, I wrote (it)’
- (13) CollP
vaqtīke āmad, be=h=eš mī-gam
when come.PST.3SG to=him DUR-say.PRS.1SG
‘as soon as he comes, I will tell (it) him’
- (14) CollP
dīdī goftam?
see.PST.2SG say.PST.1SG
‘Have you seen (it) I told (you)? (= What did I tell you!)’

4.10. Pronoun clitics

Besides the full pronouns, there is a series of pronominal clitics in SNP, as follows:

Table 4: SNP pronominal clitics

1SG	=am	1PL	=emān	(coll. =emūn)
2SG	=at (coll. =et)	2PL	=etān	(coll. =etūn)
3SG	=aš (coll./ENP(!) =eš)	3PL	=ešān	(coll. =ešūn)

When they follow a word that ends in vowel, the pronominal clitics may take a glide, or lose their initial vowel, to avoid a hiatus, see §3.9. Throughout the history of Persian, the pronoun clitics may be suffixed to verbs, prepositions and nouns, and serve the same grammatical functions as the full pronouns, with the exception of sentence subject (but see below), e. g.:

direct object

- (15) *mibīnam=et*
see.PRS.1SG=2SG
'I see you'

indirect object

- (16) *nešān=et dādam* (or, less often: *nešān dādam=et*)
sign=2SG give.PST.1SG
'I showed (it) you'

prepositional complement

- (17) *be=h=et goftam*
to=2SG say.PST.1SG
'I told (it) to you'

possessive complement

- (18) *barādar=at*
brother=2SG
'your brother'

A pronoun clitic may even be infix between a preverb and a verbal stem, e. g. *bar-aš dār!* 'remove it!' (<*bar-dāštan* 'remove'). It is also worth noting the conditions where clitic pronoun and full pronoun are not interchangeable, e. g.:

- (19) *ketāb=am=rā na-dīd-am*
book=1SG=OBJ NEG-see.PST-1SG
'I have not seen my book'

but not:

- (20) **ketāb-e man=rā na-dīd-am*
 book-EZ I=OBJ NEG-see.PST-1SG
- (21) *az īn čīz xoš=am mi-y-āyad*
 from this thing good=1SG DUR-GLIDE-come.PRS.3SG
 ‘I like this thing’

but not:

- (22) **az īn čīz xoš-e man mi-y-āyad.*
 from this thing good-EZ I DUR-GLIDE-come.PRS.3SG

In colloquial Persian, a third person singular sentence subject may be indicated by a personal clitic, e. g. *goft=eš* [say.PST=3SG] ‘he said’, *āmad=eš* [come.PST=3SG] ‘he came’. The same construction occurs already in ENP (Lazard 1963: §319–321, e. g. *gufst=aš* ‘he said’, *bud=aš* ‘he was’). It is a survival from regular MP ergative marking, although the MP personal clitic was usually attached not to a verb, but to the sentence-initial word, e. g. *u=š guft* ‘and he said’ (Durkin-Meisterernst 2014: 292). This construction may always have existed, throughout the history of Persian, as a possible, albeit (until today) non-preferred, sub-standard construction.

Combinations of more than one personal clitic are not possible in NP, i. e. in ditransitive constructions like **dādam=ešān=at* ‘I gave them to you’ (/ ‘I gave you to them’). Instead, one of them must be expressed separately (*be-h=at dādam=ešān*).

In modern SNP, the preference for a clitic direct or indirect object depends largely on the verb. E. g., with *dādan* “give”, both clitics are possible (i. e., *dādam=eš* can mean both: “I gave it” and, although less familiar, “I gave (it) to him”). As for *goftan* “say”, however, forms like *goftam=et* “I told (it) you” are theoretically possible but hardly used; *be=h=et goftam* is clearly preferred. Forms like *goftam=at* are, however, frequent in Classical Persian literature.

It is interesting to note that a 1SG pronoun clitic can be attached to a 3SG past verbal form to yield an expression that is identical with a inflected 1SG verbal form, e. g. *šenāxt=am* [recognize.PST.3SG=me] ‘he recognized me’ vs. *šenāxtam* [recognize.PST.1SG] ‘I recognized (it)’.

4.11. Demonstratives

The SNP demonstrative pronouns are *īn* ‘this’ and *ān* ‘that’, pl. *īnhā, ānhā* (in CP, there was also *īnān, ānān*). When they qualify nouns as demonstrative adjectives, they precede the noun and remain always in the sg., e. g.: *īn mard* ‘this man’, *ān mardhā* ‘those men’.

4.12. Reflexivity

4.12.1. Modern Persian

In modern Persian, reflexivity is expressed through an uninflecting reflexive marker SNP *xod* (Tajik *xud*), to which clitic pronouns and, if necessary, the object clitic *=rā* may attach:

- (23) *man xod=am=rā mī-bīnam*
 I REFL=1SG=OBJ DUR-SEE.PRS.1SG
 ‘I see myself’
- (24) *Ahmad bā xod=aš hārf mī-zanad*
 Ahmad with REFL=3SG word DUR-strike.3SG
 ‘Ahmad talks to himself’

The use of the reflexive marker is not normal, however, when a possessor is expressed, unless there is a strong sense of emphasis (see below):

- (25) *mā dūst-hā-y=emān=rā* (no *xod*!) *mī-bīnīm*
 we friend-PL-GLIDE=1PL=OBJ DUR-SEE.PRS.1PL
 ‘We see our friends’

Xod may also be used as an identity marker together with the enclitic pronoun in order to emphasize the identity of the person concerned or to contrast with other possible interpretations, e. g.:

- (26) *mā dūst-hā-y-e xod=emān=rā mī-bīnīm*
 we friend-PL-GLIDE-EZ EMPHAT=1PL=OBJ DUR-SEE.PRS.1PL
 ‘We see our own friends’
- (27) *Ahmad xod=aš mī-y-āyad*
 Ahmad EMPHAT=3SG DUR-GLIDE-COME.PRS.3SG
 ‘Ahmad himself comes’

4.12.2. Diachronic developments

In the history of NP, there was both a morphological and a systemic change in the expression of reflexivity. ENP had the two reflexive markers *xwēš* and *xwēštan* (the former used mainly for reflexive possession, the latter for all other reflexive relations). *Xwad* (later *xod*) was originally an emphatic adverb ‘even, just’, but started replacing *xwēš* and *xwēštan* in their reflexive meanings from late MP / ENP onwards, in a gradual development that took centuries to be completed (Paul 2002: 29–30).

In contrast to modern SNP *xod*, the ENP and CP reflexive markers *xwēš*, *xwēštan* and (increasingly) *xwad* were still used without additional enclitic pro-

nouns, hence no marking of person and number; and they were still necessary to mark reflexive possession, e. g.:

- (28) ENP/CP

<i>Aḥmad</i>	<i>xwēštan(=rā)</i>	<i>mī-bīnad</i>
Aḥmad	REFL(=OBJ)	DUR-SEE.PRS.3SG
'Ahmad sees himself'		

- (29) *Aḥmad bā xwad harf mī-zanad*
 Ahmad with REFL word DUR-STRIKE.PRS.3SG
 'Ahmad talks to himself'

- (30) *mā dūstān-e xwēš(=rā) mī-bīnīm*
 we friends-EZ REFL(=OBJ) DUR-SEE.PRS.1PL
 'We see our friends'

Constructions like *X. xod=rā mī-bīnad* 'X. sees himself', without personal clitic, are possible also in modern SNP, but they would represent a formal register of SNP, one that is actually emulating CP.

5. Adpositions

5.1. Prepositions

SNP Adpositions fall into four categories: simple prepositions (used without an ezafe); a large set of prepositions that are linked to their complement through an ezafe; prepositional expressions, i. e. combinations of adverbs or adjectives with simple prepositions; the postposition, or clitic, =*rā* (on which, see §5.5. below); and circumpositions (§5.6.).

5.1.1. Simple prepositions

The main SNP 'simple' prepositions, that are used without the ezafe, are the following: *az* 'from' (Abl.), *bā* 'with' (Instr., Comit.), *bar* 'upon', *be* 'to, towards' (indir. Obj./Dat.), *bī* 'without', *dar* 'in' (Loc.), *ellā* 'except', *joz* 'except', *tā* 'until, to'. Of these, *az*, *bā*, *be* and *dar* cover core case functions that are here abbreviated with the traditional terms of ablative, instrumental, etc.

5.1.2. Prepositions with ezafe

These may be divided into two groups: a smaller group of prepositions with a non-local meaning, e. g. *barā-ye* 'for (the sake of)' (Benefactive), *bedūn-e* 'without', *darbāre-ye* 'about', *mesl-e* 'like', and a larger (in principle open) class with a local

meaning, most of which go back to nouns or adverbs, e. g.: *bālā-ye* ‘over, above’, *beyn-e* ‘between’, *dāxel-e* ‘inside’, *pīš-e* ‘to(wards)’ (someone, +animate), *pošt-e* ‘behind’, *zīr-e* ‘under’. Most of these prepositions were formerly combined, as nouns, with a simple preposition, and can still be used in such a way in formal language; the simple preposition can then serve to distinguish various meanings, etc.: *dar dāxel-e* ‘inside’ (position), *be dāxel-e* ‘inside’ (direction), from the noun *dāxel* ‘interior, inside’. In less formal language, the simple prepositions *be* and *dar* can be dropped in such expressions (see also §5.4.).

5.1.3 Prepositional expressions

Prepositional expressions include examples like *ba 'd az* ‘after’, *qabl az* ‘before’, *rāje* ‘be’ ‘concerning’, or of nouns with preceding simple prepositions (containing also the “formal” varieties of the prepositions) and an ezafe, e. g.: *bā vojūd-e* ‘in spite of’ (lit., ‘with the existence of’), *be vasīle-ye* ‘by means of’, *az dast-e* ‘by (means of)’, *rū-be-rū-ye* ‘opposite’.

5.2. Historical development

The form and function of some of these prepositions changed considerably in the history of Persian. In ENP, there were historical forms like *abar*, *abā*, *bā(z)*, *ba(d)*, etc. (later > *bar*, *bā*, *be(d)*). In EJP, there was still *pa(d)* (~ENP *ba(d)*, MP *pad*), and a continuation of the MP preposition *ō* ‘to(wards)’ (pronounced [u], or [a], in EJP). The semantic shifts that some prepositions underwent are too complex to be dealt with here in detail, but it is worth mentioning that *bā* was also used for addressees of a verb of speech (e. g., *bā ō gust* ‘he said to him’), and in EJP, the preposition *u* (later *a*) was also used as a direct object marker (see Paul 2013a: §180). See the following overview, which includes also *=rā* (§5.5.):

Table 5: Historical development of prepositions

	SNP	ENP/EJP
recipient of a verb of transfer	<i>be</i>	(<i>a</i>) <i>bā</i> , <i>ba(d)</i> , <i>be</i> , <i>u</i> , <i>=rā</i>
addressee of a verb of speech	<i>be</i>	(<i>a</i>) <i>bā</i> , <i>be</i> , <i>=rā</i>
instrument	<i>bā</i> , <i>be</i>	<i>abā</i> , <i>bā(z)</i> , <i>ba(d)/pa(d)</i>
comitative	<i>bā</i>	<i>abā</i> , <i>bā(z)</i>
goal/direction of a motion verb	<i>be</i>	<i>bā(z)</i> , <i>ba(d)/pa(d)</i> , <i>be</i> , <i>u</i>
direct object of a transitive verb	<i>=rā</i>	<i>=rā</i> , <i>u</i>

5.3. Modern colloquial Persian

In modern colloquial Persian, there is a tendency to replace some of the simple prepositions by synonymous ones that are used with the ezafe, e. g. SNP *dar* ‘in’ > CollP *tū-ye*, *bar* ‘upon’ > *rū-ye* (lit., ‘[on] the surface of’), *bī* ‘without’ > *bedūn-e* (< Ar. *bi-dūni* ‘without’).

In fast or substandard colloquial speech, the ezafe may be dropped, e. g.: SNP *dar xāne* ‘in the house’ > CollP *tū-ye xāne* > *tū-xūne*. The preposition *bar* already seems to have been completely replaced by *rū-ye* in CollP; *bī* has largely fallen out of use as a (free) preposition and is mostly used as a bound prefix in adjectives like *bī=kār* ‘jobless’ nowadays. Similarly, *be* in its instrumental use has undergone lexicalisation and is used mostly in adverbial expressions like *be=zūdī* ‘soon’, *be=rāhatī* ‘simply’.

For other prepositions (simple and with ezafe), there is a similar but weaker tendency in colloquial speech to be replaced by nouns with ezafe, e. g. *be* ‘towards’ > *taraf-e* ([to] the side of’) or > *dast-e* ([to] the hand [= side] of’); *barā-ye* ‘for’ > *vāse-ye* (in fast speech, without ezafe: > *vāse*). All these prepositions belong then to the second group of §5.1.

5.4. Omission of prepositions

Another non-standard option of colloquial Persian worth of mention is the expression of location, direction or indirect objects without the prepositions *dar* or *be*, e. g.:

- (31) *man Tehrān-am*

I Tehran-COP.1SG
'I am (in) Tehran'

- (32) *raftam Tehrān* (less often, but also possible: *Tehrān raftam*)

go.PST.1SG Tehran
'I went to Tehran'

- (33) *be-deh 'Alī!*

SBJV-give.IMPER 'Alī
'give (it) to Ali!'

The same constructions are possible in Tajik dialects (*šahr raft* ‘he went to the city’, Lazard 1956: 141), and already in ENP (Lazard 1963, §713–714). This suggests that the construction has been possible in colloquial registers of New Persian throughout its history.

5.5. The particle =rā

5.5.1. Historical developments

This postposition is the most intensively discussed issue of Persian grammar. It evolved from the MP postposition *rāy* with benefactive/recipient (indirect object) sense, which was preserved in ENP =rā, e. g.:

- (34) *xalq=rā rāh-e rāst be-namāyam*
 people=rā way right SBJV-show.PRS.1SG
 ‘(that) I show the people the right way’ (Siyāsatnāme, 11th cent. C. E.)

Already in ENP, the semantic range of =rā was extended to mark also direct objects (mostly animate ones), e. g.:

- (35) *ū=rā be man sepord*
 he=rā to I entrust.PST.3SG
 ‘he entrusted him to me’ (Čahār maqāle, 12th cent.; references: see Paul 2008: 334)

During the history of NP, the functions of =rā shifted gradually away from indirect, to mark predominantly direct objects; unlike ENP, not animate but definite ones. The historical evolution of =rā, its grammaticalization from a benefactive/indirect object marker to a definite direct object marker, has been discussed in Lazard (1970). Paul (2008) has drawn attention to the importance of animacy for this process in ENP, which still plays a minor role also in SNP (Lazard 1992, §173: 185).

Although marking a definite direct object is the commonly accepted SNP core function of =rā (or coll. =ro, =o, Tajik =ro, =a), the exact nature of the conditioning factors determining the presence of =rā remains a matter of ongoing controversy. The most comprehensive and reliable study of SNP =rā continues to be Lazard (1982; see also Lazard 1992: §§173–177). The following sketch aims to present some basic constructions, and to discuss a selection of special uses.

5.5.2. Basic constructions

A SNP direct object is clearly definite (in the sense of identifiable) and takes always =rā, if it is a name or pronoun, or a noun that is complemented by a name or a (full or enclitic) pronoun, e. g.:

- (36) *Aḥmad=rā mī-bīnam*
 Aḥmad=rā DUR-see.PRS.1SG
 ‘I see Ahmad’

- (37) *māšīn-am=rā mī-forūšam*
 car-my=rā DUR-sell.PRS.1SG
 'I sell my car'

In all other cases, the presence of =rā depends on the context, and on the interplay of a number of complex factors. See the following basic examples from SNP, where =rā marks a direct object that is definite from the context, but not a generic one, or one that is marked as an indefinite object by -ī:

- (38) *ketāb=rā mī-xaram*
 book=rā DUR-buy.PRS.1SG
 'I buy the book (definite)'
- (39) *ketāb mī-xaram*
 book DUR-buy.PRS.1SG
 'I buy a book/books (generic)'
- (40) *ketāb-ī mī-xaram*
 book-INDEF DUR-buy.PRS.1SG
 'I buy a book, some book (one, indefinite)'

=rā may, however, be attached to a noun that is marked by the indefinite-specific -ī, e. g.:

- (41) *ketāb-ī=rā mī-xaram*
 book-INDEF=rā DUR-buy.PRS.1SG
 'I buy a (certain) book (specific, explained afterwards)'

The expressions *ketāb=rā* and *ketāb-ī=rā* are complementary with respect to time: while *ketāb=rā* points to the past (the book has been known beforehand), *ketāb-ī=rā* points to the future (the book will be explained soon, e. g. through a relative clause). Definition is obviously not a semantic feature of a noun, but a value that is attributed to it by the speaker, his knowledge of the context, and the situation.

The numeral *yek* 'one' may emphasize or replace the indefinite -ī in all the cases above (e. g., *yek ketābī-rā*, etc., see §4.5.1.).

5.5.3. =rā with generic objects

Just like nouns with -ī, nouns used in a generic sense are, by definition, not definite. Nevertheless, a generic object may also take =rā under certain conditions, e. g. when it is set up in contrastive opposition to another sentence constituent, and/or when it is separated from the predicate by other constituents, cf. (Lazard 1989: 280):

- (42) *xargūš dūst dāram*
 hare liked have.PRS.1SG
 'I like (to eat) hare'

- (43) *xargūš=rā hame gūštxʷārān ne-mī-xorand*
 hare=rā all carnivores NEG-DUR-east.PRS.3PL
 ‘Not all carnivores eat hare’

Sporadically a generic noun that is not emphasized/thematicized may take the =*rā*:

- (44) *serke šīr=rā mī-borrād*
 vinegar milk=rā DUR-cut.PRS.3SG
 ‘Vinegar makes milk curdle’

5.5.4. Topicalization

Lazard (1982) had already shown that besides marking a definite direct object, =*rā* can be used also to “topicalize” an important “pole” of a sentence. Recent studies like Dabir Moghaddam (1992) or Meunier and Samvelian (1997) also focused on pragmatic factors like topicalization, thematization etc. See the following example of topicalisation by =*rā*:

- (45) *otāq=o, dar-eš=o bastam* (colloquial)
 room=rā door-his=rā close.PST.1SG
 ‘as for the room, I closed its door’ (Dabir-Moqaddam 1992: 553)

5.5.5. Special verbs with =*rā*

The question as to whether semantic features of the verb or noun have any influence on the presence of =*rā* has been rather neglected so far. Lazard (1982: 186) noticed that verbs with low semantic content are less likely to occur with =*rā*, but he did not develop this argument. Indeed, there is a group of transitive verbs, semantically defined as “verbs that exert a clear and lasting influence on the direct object”. These verbs practically cannot be used without =*rā* in natural language use, even if a direct object is marked as indefinite (and not specific) by -*ī*. This includes verbs like: *koštan* ‘kill’, *tayyīr dādan* ‘change’, *xarāb kardan* ‘destroy’, *bombārān kardan* ‘bombard’. See the example:²

- (46) *man (yek) mard-ī=rā koštam*
 I (one) man-INDEF=rā kill.PST.1SG
 ‘I killed a man’ (not specific)

A use of *koštan* without =*rā* would be feasible with a generic understanding of the object, but this sounds odd and unnatural:

² I owe this important point to a discussion with Maximilian Kinzler and Ramin Shaghaghi.

- (47) *emrūz man mī-ravam mard mī-košam (?)
 today I DUR-go.PRS.1SG man DUR-kill.PRS.1SG
 ‘Today, I go killing men (generically, i. e., like a profession)’

5.5.6. Special nouns and =rā: abstract nouns

Most studies so far have concentrated on countable, concrete nouns like: *ketāb* ‘book’, *mard* ‘man’, *xāne* ‘house’, *xargūš* ‘hare’ as direct objects, sometimes also on non-countables like *āb* ‘water’ or *gūšt* ‘meat’. Abstract nouns like *āzādī* ‘freedom’ have hardly been considered. It is difficult to say if, e. g., ‘freedom’ as such, is definite or not. In simple sentences, abstract nouns seem to be used without =rā, e. g.:

- (48) rāhatī / āzādī be dast āvardam
 rest / freedom to hand bring.PST.1SG
 ‘I attained rest / freedom (etc.)’

But like generic nouns, they may do so when they are in contrastive focus (=rā is optional here):

- (49) āzādī=rā dūst dāram, na bandegī=rā
 freedom=rā friend have.PRS.1SG not slavery=rā
 ‘I love freedom, not slavery’

Abstract nouns that tend to take =rā when used as direct objects include *zendegī* ‘life’ or *marg* ‘death’, which usually refer to the speaker and is therefore “inherently definite” in sentences like:

- (50) zendegī=rā dūst dāram
 life=rā friend have.PRS.1SG
 ‘I love life’

5.5.7. Sciences, techniques, etc. with =rā

Another special group of nouns are sciences, techniques, arts and languages. They occur usually without =rā:

- (51) Fārsī balad-am
 Persian informed-COP.1SG
 ‘I know Persian’
- (52) riyāzīyāt / tārīx (etc.) tadrīs mī-konam
 mathematics / history teaching DUR-do.PRS.1SG
 ‘I teach mathematics / history (etc.)’,

But they may take $=rā$ when they are in contrastive opposition:

- (53) *emrūz Fārsī=rā tadrīs mī-konam, fardā adabiyāt=rā*
 today Persian= $rā$ teaching DUR-do.PRS.1SG tomorrow
literature=rā
 ‘Today I teach Persian, tomorrow literature’

A science that takes a nominal complement with the ezafe, becomes definite and takes $=rā$:

- (54) *tārīx-e Āīn=rā dūst dāram*
 history-EZ China= $rā$ friend have.PRS.1SG
 ‘I love Chinese history (lit., the history of China)’,

unless the whole noun phrase corresponds to the name of a subject that is regularly taught at academic institutions:

- (55) *tārīx-e Āīn tadrīs mī-konam*
 history-EZ China teaching DUR-do.PRS.1SG
 ‘I teach (the subject:) “Chinese history”’

In this example, *tārīx-e Āīn* represents the conventional name of an academic subject. The exceptional absence of $=rā$ in such cases may be due to the verb *tadrīs kardan*, which occurs typically with names of academic subjects.

5.6. Circumpositions

Besides the prepositions and the postposition $=rā$, a small number of circumpositions are attested in ENP, and in some CP texts, though no longer in use in SNP. They can be divided into two groups (other West Iranian languages such as Kurdish continue to make extensive use of circumpositions, see Haig, this volume, chapters 2.3 and 3.3):

- Combinations of (mostly simple) prepositions with $=rā$, see Lazard (1963: §539–552), e. g. *ba ... =ra* ‘to, towards’, *az ... =rā* ‘because of’, in EJP also ‘concerning’ (Paul 2013a: §184.a); *azmar-i ... =rā* ‘for’ occurs only in EJP (Paul 2013a); *mar ... =rā*, which has the same meaning as plain $=rā$, is a special case, because *mar* without $=rā$ is not a well-established preposition of ENP (Lazard 1963: §752–756).
- Combinations of various simple prepositions with each other. In ENP, it was quite common for *ba* to be combined with other simple prepositions in postpositional use, e. g. *ba ... (an)dar* ‘in’ (Lazard 1963: §619–621), or a “doubling” of prepositions like *(an)dar ... (an)dar* ‘in’ (Lazard 1963, §623).

6. Verbal morphology

6.1. Generalities

Verbal forms of all varieties of NP are based on two stems, present and past, from which a system of tenses and moods is built with the help of endings, verbal particles, and auxiliaries. The tense systems of the modern varieties of Persian, especially the written standard forms, are structurally similar and share certain parallel developments, but they are quite different formally. Tajik dialects have developed modal categories such as “presumptive” or “dubitative” that do not exist in Fārsī.

In ENP the basic verbal forms of later SNP and Tajik were already in use, but the various verbal particles (see below) became fully grammaticalized only in the course of later centuries. The tense system of ENP was thus structurally very different from SNP. This was also because it was less standardized and allowed for many regional varieties. A recent account of the evolution of the NP verbal system has been given by Lenepveu-Hotz (2014).

6.2. Certain verb classes

6.2.1. Compound verbs

Certain NP simple verbs (e. g., *zadan* ‘strike’) can combine with nouns or adjectives (e. g., *harf* ‘letter, speech’) to form “light verb constructions”. These constructions, which are very widespread in New Persian and show various levels of amalgamation, correspond to simple verbs in many other languages (*harf zadan* ‘speak’, lit. ‘to strike a word/speech’). A correct understanding of these is essential for a correct understanding of the syntax of the Persian verb phrase (see Ahadi 2001 and Samvelian 2012).

There is no clear-cut distinction between ‘light verb constructions’ and ‘phrasal verbs’ with a more general meaning (e. g., *māhī gereftan* ‘to fish, angle’, lit. ‘to catch a fish’), as shown by various syntactic tests in Ahadi (2001: 220–228). Both groups form a continuum, on the loose edge of which follow verbal constructions with a generic noun like *ketāb xarīdan* ‘buy books’.

Other groups of compound verbs include verbs with preverbs, e. g. *bar-dāštan* ‘remove’; prepositional verbal constructions like *be donyā āmadan* ‘be born’ (lit., ‘come to the world’); and verbal constructions with ezafe like: *montazer-e ... būdan* ‘wait for ...’, with a participle (*montazer* ‘waiting’) followed by the object with the *ezafe*.

6.2.2. Impersonal verbal constructions

A large and increasing number of verbal constructions in CollNP is formed with a 3SG verb and an experiencer, often a personal clitic, to express physical and psychological processes, e. g.:

- (56) *xoš-am mī-y-āyad*
good-me DUR-glide-come.PRS.3SG
'I like (it)'

- (57) *xʷāb-at bord*
sleep-you bring.PST.3SG
'You fell asleep'

- (58) *gorosne-am ast*
hungry-me COP.3SG
'I am hungry'

The latter is used alongside the SNP personal construction *gorosne-am* 'I am hungry' (see Rāsex Mahand 2007).

6.2.3. Auxiliary verbs

NP uses the following auxiliary verbs: *xʷāstan* 'want' (for the future, §6.7.), the copula (compound past tenses, §6.8.), and *šodan* 'become' (passive, §6.13.). For the progressive, SNP uses *dāstan* 'want', *Darī raftan* 'go', and Tajik *istodan* 'stay' (§6.12.).

6.3. Person and number agreement suffixes

These are almost identical in each variety of NP, and across all tenses. The sole exception is SNP 3SG, which is *-ad* in the present, and *-Ø* in the past tenses. The enclitic present copula is identical with these endings, with the exception of 3SG (§6.5.). The endings in overview, with some of the more important regional/chronological variations are as follows:

Table 6: Person and number agreement in Persian

	SG	PL
1	<i>-am</i> (ENP/Taj./Darī also <i>-om</i>)	<i>-īm</i> (ENP <i>-ēm</i> , Taj. <i>-em</i>)
2	<i>-ī</i>	<i>-īd</i> (ENP <i>-ēd</i> , Taj. <i>-e(d)</i> ; CollP <i>-īn</i> , Darī <i>-ēn</i>)
3	<i>-ad</i> (coll. <i>-e</i>) (present)/- Ø (past)	<i>-and</i> (coll. <i>-an</i>)

Politeness is expressed with the 2nd person plural. For impersonal 'one', the 3rd person plural is used, e. g. *mī-gūyānd* [DUR-say.PRS.3PL] 'they say, one says'.

6.4. Verbal affixes

The following affixes (Table 7) may be combined with Persian verbs to express tense, modality, and negation:

Table 7: Verbal affixes

<i>mī</i> (ENP (<i>ha</i>) <i>mē</i> , Darī <i>mē</i> , Taj. <i>me</i>):	indicative present tense, imperfect (used with all verbs except <i>dāštan</i> ‘have’)
<i>be</i> (allomorphs <i>bo</i> , <i>bī</i> etc. depending on verbal stem; ENP <i>bi</i> or <i>bu</i>):	present subjunctive and imperative
<i>na</i> (SNP <i>ne</i> before <i>mī</i>):	negation
<i>ma</i> (= ENP, in modern Persian <i>na</i>):	prohibitive
- <i>ē</i> (ENP - <i>ē</i> (<i>ð</i>), CP - <i>ē</i>):	irrealis/habitual of the past (no longer used in modern Persian)

Negative *na-* can be combined with *mī-/me-* (always preceding it, e. g. SNP *ne-mī-ravam* ‘I don’t go’), but not with *be-*, which it replaces. In ENP, many particle combinations were still possible that fell out of use later, e. g. (*ha*)*mē* + *bi-* (Lazard 1963: §357) or *bi-* + *na-* (Lazard 1963, §443), e. g. *bi-na-y-āmad* [SBJV-NEG-GLIDE-come.PRS.3SG] ‘He did not come’. For verbs with preverbs, in SNP a prefix is always inserted between the preverb and the verbal stem, while in Tajik and (partly) Darī (e. g., Kabuli), it regularly precedes the prefix, e. g.:

(59) SNP

bar-mī-āyad
out-DUR-come.PRS.3SG
'he comes out, appears'

(60) Taj.

me-bar-oyad
DUR-out-come.PRS.3SG
'he comes out, appears'

6.5. Present tense, imperative

The SNP/Taj. present tense is formed by *mī-/me-* + the verbal present stem + the endings of Table 6, e. g. present indicative SNP *mī-rav-am* ‘I go’. The present subjunctive is indicated by SNP *be-*, which may be dropped in certain cases (e. g., with compound verbs). In central and northern Tajik dialects, the subjunctive is expressed by a bare present stem, e. g. *kun-am* ‘(that) I may do’ (Lazard 1956: 145). In some Tajik verbs, a historical *be-* (or *bu-* etc.) has been lexicalized and integrated into the verbal stem, e. g. *buraftan* ‘go’.

The present tense also serves to express the future in SNP and CollNP, e. g.:

- (61) *fardā man mī-ravam bāzār*
 tomorrow I DUR-go.PRS.1SG bazaar
 ‘Tomorrow, I shall go to the bazaar’

The imperative is expressed by *be-/bo-* (etc.) and the bare present stem for the sg., and with an additional 2PL ending for the plural, e. g.: *bo-xor* ‘eat!’ / *bo-xorid* ‘eat (PL)!’

In CollNP (Fārsī and Tajik), a small number of verbs that are of high frequency exhibit shortened forms of the present stem, e. g.: *mī-r-am* ‘I go’, *mī-g-am* ‘I say’, *mī-d-am* ‘I give’ (SNP *mī-rav-am*, *mī-gū-y-am*, *mī-dah-am*; Tajik coll. *mē-r-am*, etc.).

6.6. The copula

There are two forms of the SNP present tense copula, a full and an enclitic one:

Table 8: Two forms of the SNP present tense copula

‘to be; to be there’ (full)		‘to be’ (enclitic)
1SG	<i>hastam</i>	-am
2SG	<i>hastī</i>	-ī
3SG	(<i>h</i>) <i>ast</i>	-ast (coll. -e; after long vowel: -st, coll. -s)
1PL	<i>hastīm</i>	-īm
2PL	<i>hastīd</i>	-īd (coll. -īn, Darī -ēn)
3PL	<i>hastand</i>	-and (coll. -an)

The full form of the copula can be used with a stress on the verbal stem in the emphatic sense of: *man hástam* ‘I am here’, but also as a normal substantive verb, like the enclitic form, cf.

- (62) *man Ālmānī-y-am* / *man Ālmānī hastam*
 I German-GLIDE-COP.1SG / I German COP.1SG
 ‘I am German’.

The copula is necessary in sentences stating identity or quality. It is not used in exclamations and other elliptic expressions, e. g.:

- (63) *xodā=rā šokr*
 God=rā thanks
 ‘Thanks (be) to God’ (also in ENP, Lazard 1963: §779)

- (64) *xoš be hāl-at*
 good for condition-your
 ‘Good for you!’

and with *kū* ‘where?’:

- (65) *Hasan kū?*
 Hasan where?
 ‘Where is Hasan?’

The copula has the simple past *būd-* (from the infinitive *būdan* ‘be’) and the irregular subjunctive stem *bāš-*, e. g. the imperative sg.: *movāzeb bāš!* ‘be attentive / take care!'

The negative form of the copula is: *nīst-am* ‘I am not’, *nīst-ī* ‘you are not’, etc.

6.7. Future

Besides the present tense, which may express the future, especially in CollNP (§6.4.), there is a morphological future tense formed with the present stem of the verb ‘to want’, the appropriate personal endings and the shortened infinitive, e. g.: *xʷāham raft* ‘I shall go’.

This tense did not yet exist in ENP, in which *xʷāham raft* would still mean ‘I want to go’ (also expressed by *xʷāham raftan* with the full infinitive). Of course the development of future meanings from a desiderative ‘want’ is a well-attested grammaticalization pathway, but it was only with the development of modal constructions with the subjunctive in CP (§6.15.) that the future developed as a separate tense.

6.8. Past tenses

The basic system of past tenses of SNP and literary Tajik is based on three forms: simple narrative past (*raft* ‘he went’), durative-iterative imperfect (*mī-raft* / *me-raft* ‘he used to go’), both with the appropriate personal endings of the past (§6.2.); and present perfect, built with the past participle (*rafe* / *rafta* ‘gone’) and the enclitic copula (§6.6.): *rafte-ast* / *rafta-ast* ‘he has gone’.

In all varieties of NP a combination of imperfect and present perfect has also developed, expressing in SNP a durative past action, which is completed in the past, e. g.:

- (66) *bīš az emrūz az xodā mī-tarsīde-and*
 more than today from God DUR-fear.PTCP-COP.3PL
 ‘(... a time when the poor) were more in fear of God than today’ (Lazard 1992: 154–155)

In Tajik this verbal form developed into a resultative mode (Lazard 1956: 151). It is already attested, though sparsely, in ENP (Lazard 1963: §389).

The Persian tense system is completed by the pluperfect that is built from the PPP and the past copula (*rafte būd* ‘he had gone’).

6.9. Present perfect

In colloquial registers of both Fārsī and Tajik, the present perfect endings are amalgamated with the stem, and can partly be distinguished from the simple past only by stress (e. g., Fārsī *rafté-am* ‘I have gone’ > *raftá(a)m*: simple past *ráftam* ‘I went’; Tajik *raftá-y-am* ‘I have gone’ > *raftém*: simple past *ráftam*, Lazard 1956: 147). In CollNP, the 3SG present perfect has lost its ending (*rafté* ‘he has gone’), which can be either explained through the colloquial form of the copula (*e* instead of *ast*, thus: *rafté-e* > *rafté*), or by analogy with the zero form of 3SG in the simple past (cf. Table 6).

6.10. Tajik “resultative”

In Tajik dialects, on the basis of the present perfect a new “resultative” verbal mode has emerged (Lazard 1956: 148–151). On closer inspection, Tajik dialects offer a lot of constructions that do not exist in Fārsī. For example, derivations from a past participle in *-agī* (not found in Fārsī) can be used both as nouns (*omadagī-am* ‘my coming’) and verbally, in constructions like: *az man omadagī* ‘since I have come’ (Lazard 1956: 171). This and other constructions show a tendency in Tajik towards clause subordination through non-finite verb forms (participles), as opposed to the predominance of finite verbs used in subordination in Fārsī.

In the Tajik dialect of Varzob, the participle in *-agī*, together with the copula, forms a separate verbal mode, the “presumptive”, e. g.:

- | | | | |
|------|-------------------------|---|--|
| (67) | <i>qošuq kujo?</i> | — | <i>xona budagī-st</i> |
| | spoon where? | — | house be.PST.PTCP-COP.3SG |
| | ‘where are the spoons?’ | — | they must be (/ are presumably) in the house.’ |

A presumptive exists also in the Darī dialect of Kabul, but it is built in a different way, with the adverb *xāt* (< *xʷāhad* ‘he wants’, Lazard 1956: 160). The presumptive can be considered a sub-type of evidential, or indirect source. Turkic (Uzbek) influence may have played a role in this development.

Forms in *-agī* are not found in ENP in Arabic script, but they occur in certain EJP exegetical texts (*tafsir*) (Paul 2013: §161), e. g. *kardagī* [correct for: *kardageē*] ‘he has done’, often with a resultative meaning.

6.11. ENP

The verbal system of ENP is difficult to describe, since it is not uniform or standardized and has a lot of regional/dialectal varieties. The earliest stage of Standard Persian, the language of the classical authors of the 13th–15th centuries, represents a transitory stage of the language, in which many ENP grammatical features are still found, dialectal standardisation has been attained, and the rules of modern Standard Persian are gradually evolving. Some important features in which the ENP verbal system diverges from the later SNP one include:

- a) In some of the earliest EJP documents, the present perfect with the past participle in *-a* (SNP *-e*) had not yet been fully developed as a tense, participles like *nibišta* had still an adjectival-static (passive) meaning, e. g., in expressions like: *nibišta hest* ‘it is written’ (Paul 2013: §165.b). In most ENP texts, the present perfect with *-a* had already been fully developed, but there are still examples of passive-adjectival past participles (e. g., *man kušta am* ‘I am killed’, Lazard 1963: §487; in modern SNP examples of lexicalized past participle, e. g. *baste* ‘closed, tied; package’).
 - b) The EJP present perfect can still be built with the plain past stem, e. g., *nibišt hest* ‘he has written’ (Paul 2013a: §164.b). In ENP, this is very rare and is practically restricted to one manuscript (Lazard 1963: §485).
 - c) In ENP, there is a variety of auxiliaries used for compound past tenses that no longer exist in SNP, like *hast-*, *buv-*, *b-*, etc. (e. g., *dīda hastē* ‘have you (PL) seen ...?’, Lazard 1963: §482–483).
 - d) In some ENP texts from northeastern Iran (e. g., in the national epic *Šāhnāme*) there is the “present perfect II” of the *kardastam* (1SG), *kardastī* (2SG) etc. type, probably of dialectal origin, which is used besides the “normal” present perfect, in the same sense. It occurs occasionally also in later Classical poetry, but has vanished from SNP.
 - e) (*ha)mē* and *bi-* were still prosodically independent particles in ENP and would be grammaticalized only in the course of Classical literature. In early CP, *bi-* was often used also with the simple past (e. g., *bi-raft* ‘he went’).
 - f) In ENP and (early) CP (but no longer in SNP) there was a suffix *-ē(δ)* (EJP *-ē(h)*), deriving historically from a MP optative particle, which had the meanings of irrealis (conditional, potential, etc.) and habitual, e. g. (conditional):
- (68) *agar man ānjā būdam-ē*
 if I there be.PST.1SG-IRR
 ‘if I had been there’ (Lazard 1963: §456)

6.12. Progressive

Both Fārsī, Kabuli Darī and Tajik have developed a progressive verbal mode, present and past, with a similar function but a completely different structure, see Table 9 (Lazard 1956: 160, 165–166):

Table 9: Progressive verbal mode

	Present	Past
SNP	<i>dāram mī-ravam</i> ‘I am going’	<i>dāštam mī-raftam</i>
Kabuli	<i>rafta mē-ravom</i>	<i>rafta mē-raftom</i>
Tajik	<i>rafta istodam</i> (in dialects shortened to: <i>rafsodam</i>)	<i>rafta istoda budam</i> (to: <i>rafsodudam</i>)

It is worth noting that the progressive mode does not allow negativization, i. e. there is no **nadāram mīravam* or **dāram nemīravam*. Typologically, this may reflect a universal tendency: it does not make much sense to point to an ongoing action or process in progress, if it does not take place. Besides indicating that an action is still in progress, there is an “intentional” use of the progressive mode in CollP., referring to intended actions in the future, e. g.: *dāram fardā mī-ravam Ālmān* ‘I am about to go to Germany tomorrow’.

6.13. Passive

A SNP passive is built with the past participle and the auxiliary verb *šodan* ‘become’, e. g. *xarīde šod* ‘(it) was bought’. The passive is mainly employed when the agent is unknown; in formal language an agent may be added with compound prepositions like *az taraf-e ...* ‘by, from ... side’. In ENP, there was also a passive with *āmadan* ‘come’, e. g. *gufta āmad* ‘it was said’.

6.14. Specific usages

There are specific, noteworthy usages of the SNP tenses that differ from that of English (Lazard 1992: §131):

A present verb may be used to express a persistent situation where English uses the present perfect, e. g.:

- (69) *čand sāl ast īnjā zendegī mī-konī?*
how.many year COP.3SG here living DUR-do.PRS.2SG
'For how many years have you been living here?'

- (70) *az zamān-e ... sī sāl mī-gożarad*
 from time-EZ ... thirty year DUR-pass.PRS.3SG
 'From the time of..., thirty years have passed.'

A past verb may indicate an action that is accomplished (and witnessed) at the very moment of speaking, e. g.:

- (71) *āmadand*
 come.PST.3PL
 'They have (just now) come'

or going to be completed in the immediate future (Lazard 1992: §138):

- (72) *āmadam!*
 come.PST.1SG
 'I am coming (just now)!'

A past verb may also be used in a future context in conditional clauses (§8.4.2.3.).

6.15. Modal verbal constructions

In the history of NP, constructions involving modal verbs changed their structure substantially, the best example being the verb *tavānistān* 'to be able'. While in ENP and CP, *tavānistān* was followed by the full or short infinitive, e. g.:

- (73) *tavānad raft(an)* (or, inverted: *raft(an) tavānad*)
 can.PRS.3SG go.INF
 'He can go'

in modern SNP it is complemented by a subjunctive verb (Paul 2002: 26–28):

- (74) *mī-tavānad be-ravad*
 DUR-can.PRS.3SG SBJV-go.PRS.3SG
 'He can go'

An intermediary stage was a dependent clausal construction with subordinating particle *ke* 'that' (*mī-tavānad ke be-ravad* 'He can go'). SNP impersonal modal constructions with the short infinitive like *bāyad raft* 'one has to go' can be regarded as remnants of the older constructions.

In Tajik dialects, and in Kaboli Darī, a construction with the passive participle has developed (*rafta mē-tavānam* 'I can go'), but what looks like a past participle in these constructions has been explained as being derived historically from an old infinitive (*raftan*) whose *-n* has been dropped (Lazard 1956: 176). The inverse order of modal and full verb here, in contrast to modern Fārsī, may reflect Turkic influence.

Table 10: Overview of the most important developments of the NP tense system

	SNP	← ENP	→ Tajik
Present:	<i>mīravam</i>	([ha]mē) <i>rawam</i>	<i>meravam</i>
Subjunctive:	<i>beravam</i>	(bi) <i>rawam</i>	<i>ravam</i> (be-: lexicalized)
Future:	<i>xʷāham raft</i>	—	<i>xoham raft</i>
Simple past:	<i>raftam</i>	(bi) <i>raftam</i>	<i>raftam</i>
Imperfect past:	<i>mīraftam</i>	([ha]mē) <i>raftam</i>	<i>meraftam</i>
Present perfect (I):	<i>rafte-am</i>	<i>rafta-am, rafta buvam etc.</i> (EJP) <i>raft hem</i>	<i>rafta-am</i>
Present perfect (II):	—	<i>raftastam</i>	—
Durative present perfect:	<i>mīrafte-am</i>	([ha]mē) <i>rafta-am (rare)</i>	<i>merafta-am (resultative)</i>
Irrealis:	—	([ha]mē/bi) <i>raftam-ē</i>	—
Resultative:	—	(EJP) <i>raftagī</i>	<i>raftagī</i>
Modal construction:	<i>mītavānad</i> <i>beravad</i>	<i>tavānad raft(an)</i>	<i>rafta metavonad</i>

7. Syntax of basic clauses

This section on syntax presents only the basic constructions, and highlights, very selectively, a small number of noteworthy phenomena. This applies especially to the syntax of complex sentences. An excellent comprehensive description of SNP syntax can be found in Lazard (1992: 177–257).

7.1. Noun phrases

In SNP, the modifier of a noun phrase, be it a substantive or descriptive adjective, always follows the head noun with an unstressed connecting particle -(y)e, which is called “ezafe”, e. g.:

- (75) *xāne-ye pedar*
house-EZ father
‘the father’s house’ (N-N)
- (76) *mard-e bozorg*
man-EZ big
‘the big man’ (N-Adj)

Demonstrative, quantifying and indefinite adjectives always precede the noun and remain in the singular, e. g., *īn xāne* ‘this house’, *čand xāne* ‘some (/ how many) houses’.

In Tajik dialects, and in the Darī of Kabul, a construction with an inverted ordering of elements has developed, e. g.:

- (77) *muallim-a kitob-aš*
 teacher-to book-his
 ‘the teacher’s book’ (Lazard 1956: 182, 184; Windfuhr and Perry 2009: 443)

This is most conspicuous in the northern Tajik dialects that have been under heavy Uzbek influence. In these Tajik dialects, adjectives also exhibit a tendency to precede the modifier without ezafe (Lazard 1956: 182). A similar construction without *-a* is widespread also in colloquial registers of modern Fārsī, e. g.:

- (78) *Ahmad zan-aš kojā-st?* (for SNP: *zan-e Ahmad kojā-st?*)
 Ahmad wife-his where-COP.PRS.3SG
 ‘Where is Ahmad’s wife?’

In ENP, the word order of noun phrases was still less strict than in SNP, and certain high-frequency adjectives like ‘big, small, good, bad’ could precede the head without the ezafe (Lazard 1963, §165–172, e. g. *bað mardān* ‘the bad men’, Lazard 1963, §167; similarly in EJP, Paul 2013a: §189.b). Less often, modifying nouns could precede the head, e. g.:

- (79) *jihān šahriyār*
 world ruler
 ‘the ruler of the world’ (Lazard 1963: §163).

In ENP and CP, an *-ē* that marked the head noun of a noun phrase as indefinite (§4.5.) replaced the ezafe before the descriptive adjective, e. g.:

- (80) *mard-ē bozorg*
 man-INDEF big
 ‘a great man’

In SNP, and in high registers of spoken Fārsī, this construction still exists as an emulation of classical style. In “normal” spoken Fārsī, the indefinite noun phrase retains the ezafe, and the *-ī* (< ENP/CP *-ē*) is attached at the end of the noun phrase (*mard-e bozorg-ī*).

7.2. Word order

7.2.1. SNP

For the word order of a simple SNP sentence, the following rules apply:

- a. The subject (S), if overtly expressed at all (see §4.8. on pro-drop), precedes the predicate (Pred). The predicate may be a full verb or a copula.
- b. Direct or indirect objects (Od, Oi), goals of verbs of motion (Dir), and predicate complements (CoPr) are inserted between S and Pred. With regard to the relative ordering of Od and Oi in SNP, see Faghiri et al (2014) for a recent corpus-based study.
- c. In ditransitive sentences, a definite Od precedes the Oi, an indefinite Od follows it.

Examples for these rules:

- (81) a. *Hasan mī-ravad* (S–Pred)
 Hasan DUR-go.PRS.3SG
 ‘Hasan goes’
- b. *Hasan 'Alī=rā mī-bīnad* (S–Od–Pred)
 Hasan 'Alī=rā DUR-see.PRS.3SG
 ‘Hasan sees Ali’
Hasan be Tehrān mī-y-āyad (S–Dir–Pred)
 Hasan to Tehran DUR-GLIDE-come.PRS.3SG
 ‘Hasan comes to Tehran’
Hasan tambal ast (S–CoPr–Pred)
 Hasan lazy COP.3SG
 ‘Hasan is lazy’
- c. *Hasan ketāb=rā be 'Alī dād* (S–Od.def–Oi–Pred)
 Hasan book=rā to 'Alī give.PST.3SG
 ‘Hasan gave the book to Ali’
Hasan be 'Alī ketāb dād (S–Oi–Od.indef–Pred)
 Hasan to 'Alī book give.PST.3SG
 Hasan gave Ali a book/books’

In SNP, exceptions to these rules may occur, especially if a constituent is emphasized. In colloquial registers, almost any exception is possible, or even usual. For example an object may follow the predicate if emphasized (Lazard 1992: §193):

- (82) *'omr-ešān=rā* (Od) *dādand* (Pred) *be šomā* (Oi)
 life-their=rā give.PST.3PL to you
 ‘They gave you their life’ (i. e., idiomatically, ‘they passed away’)

- (83) *šomā* (S) *dīdīd* (Pred) *ān* *yārū=rā* (Od)?
 you see.PST.2PL that fellow=rā
 ‘Did you see the fellow in question?’

A direct object may also be topicalized and be placed before the subject. Less often, a predicate precedes the subject when it is strongly emphasized, e. g.:

- (84) *tamām* *na-šod* (Pred) *kār-et* (S)?
 finished NEG-become.PST.3SG work-your
 ‘Isn’t your work finished?’ (Lazard 1992: §193)

Besides emphasis and topicalisation, there may be a “rhythmic” reason to break the word order rules, in that very long noun phrases in object or copula complement function tend to follow the predicate (this applies also to SNP), e. g. (Lazard 1992: §193):

- (85) *īn* (S) *būde-ast* (Pred) *natiye-ye* ‘adam-e *tavajjoh* (CoPr)
 this be.PTCP-COP.3SG result-EZ lack-EZ attention
 ‘This was the result of lack of attention’

Following Windfuhr and Perry (2009: 480), literary Tajik seems to comply with the word order rules, and exceptions, of SNP. The most frequent deviation of word order in Tajik is object fronting (OSV, OVS). For the dialects of Tajik, no study of word order is currently available.

7.2.2. ENP

In ENP, word order is said to be much freer than it is in SNP (Lazard 1963: §788, who however does not distinguish between Od and Oi). The study of EJP, which may be taken to represent ENP, has identified word order rules that are not as strict as in SNP, but already quite similar to them.

ENP exhibits a preference for S–O–Pred, but where constituents are given emphasis, many exceptions are possible, e. g. Pred–S (Lazard 1963: §796, §799), or even O–Pred–S, possibly reflecting colloquial style (Lazard 1963: §801). In EJP, Oi usually follows the Od, but can also follow the Pred, as in modern colloquial Persian. The two preferred possibilities for the main constituents’ position would be S–Od–Oi/Dir–Pred and S–Od–Pred–Oi/Dir (Paul 2013a: §193). Exceptions occur, e. g. Od follows Oi (S–Oi–Od–Pred, Paul 2013a: §199), or it follows even the Pred (Paul 2013a: §195.c):

- (86) *pa* *čē* *bi-šināsī* (Pred) *durstīh* *i-ēn* *tis* (Od)
 by what SBJV-know.PRS.2SG truth EZ-this matter
 ‘How do you know the correctness of this matter?’

The case of sentence-initial Pred in legal document should be considered separately, since this seems to be a peculiarity of legal formularies (Paul 2013a: §194.b), e. g.:

- (87) *u-guft* (Pred) *ēn* *Danyāl* (S)
 and-say.PST.3SG this Daniel
 ‘and this Daniel said: ...’

To summarize, the main differences between ENP and SNP seem to be:

in ENP, the Oi/Dir could more often follow the Pred, whose final position is less strict.

in ENP, the Od was more strictly positioned before the Oi than in SNP.

8. Syntax of complex clauses

8.1. Noun clauses

8.1.1. SNP

SNP clauses that are dependent on verbs of belief, speaking, wishing, etc., regularly follow the main clause, are introduced by the particle *ke* ‘that’, and always have a finite verb form. Direct speech is preferred. The tense of the dependent clause is independent of that of the main clause, e. g. future (dependent) following past (main clause). Indirect speech is also possible, in which case only the personal suffix changes, as seen in the following examples (taken from Lazard 1992: §203–204):

- (88) *Aḥmad goft ke pedar-am xʷāhad āmad*
 Ahmad say.PST.3SG that father-my will come.INF
 ‘Ahmad said that his father would come’ (dir. speech)
- (89) *Aḥmad goft ke pedar-aš xʷāhad āmad*
 Ahmad say.PST.3SG that father-his will come.INF
 ‘Ahmad said that his father would come’ (indir. speech)

After verbs meaning ‘want, wish, tell (to do), hope’ etc., the verb in the dependent clause is in the present subjunctive:

- (90) *meyl dāštam ke be Tehrān be-ravam*
 inclination have.PST.1SG that to Tehran SBJV-go.PRS.1SG
 ‘I wanted to go to Tehran’
- (91) *mo’allem be šāgerdhā goft ke ketābhā-y-etān=rā
 teacher to pupils said that books-GLIDE-your=your=rā
 be-gīrīd
 SBJV-take.2PL*
 ‘The teacher told the pupils to take their exercise-books’

- (92) *mo'allem be šāgerdhā goft ke ketābhā-y-ešān=rā
 teacher to pupils said that books-GLIDE-their=rā
 be-gīrand
 SBJV-take.3PL*
 ‘The teacher told the pupils to take their exercise-books’

In colloquial Fārsī the *ke* is often dropped:

- (93) *Ahmad goft pedar-am
 Ahmad say.PST.3SG father-my
 mī-y-ād
 DUR-GLIDE-come.PRS.3SG (present for future)
 ‘Ahmad says his father would come’*

- (94) *Ahmad goft mī-y-ād-aš
 Ahmad say.PST.3SG DUR-GLIDE-come-3SG (cf. §4.9.2.)
 ‘Ahmad said he would come’*

The dependent clause may be topicalized, in which case it is introduced by a cataphoric demonstrative expression, and then precedes the main clause (Windfuhr and Perry 2009: 512):

- (95) *īn (xabar) ke 'Alī xāne=rā sāxt dorūy ast
 this news that Ali house=rā build.PST.3SG lie COP.3SG
 ‘(The news) that Ali built a house is a lie’*

8.1.2. ENP

The few examples in EJP (Paul 2013a: §204) and the remarks in Lazard (1963: §810) show that also in ENP direct speech was possible in these types of sentence.

8.1.3. Tajik

In Tajik (Windfuhr and Perry 2009: 513), direct speech seems to be impossible, or unusual, in these types of sentence. Sentences of the indirect speech type (v.s.) are possible, but in literary Tajik and Tajik dialects, Uzbek-type nominal clauses that precede the main clause have become widespread, e. g. (Windfuhr and Perry 2009: 514):

- (96) *xud-i ū ki-st? gufta man az Rahim Qand
 REFL-3SG he who-COP.3SG say.PTCP I from Rahim Qand
 pursidam
 ask.PST.1SG
 ‘Who is he, actually? I asked Rahim Qand’ (literally: ‘Who is he? – [that]
 having said, I asked from Rahim Qand’)*

8.2. Relative clauses

SNP dependent relative clauses follow their head noun, which often has the “specific” *-ī* (§4.5.2.), and contain finite verb forms. They are introduced by the invariable relative particle *ke*, e. g.:

- (97) *mard-ī ke mī-bīnī dozd ast*
 man-*ī* REL DUR-see.PRS.2SG thief COP.3SG
 ‘The man you see is a thief’

The SNP relative clause has been dealt with in detail by Lehmann (1978; see also Lazard 1992: §207–213). In ENP the specifying *-ē* (> SNP *-ī*) occurred more rarely than in SNP (Lazard 1966), in EJP it occurs very rarely (Paul 2013a: §207.e). EJP has even preserved, from MPī, the ezafe *i* introducing relative clauses (besides *ki*; Paul 2013a: §206). Şādeqī (2016) has recently shown that it exists, very rarely, also in ENP.

The data in Windfuhr and Perry (2009) suggest that the SNP strategies for constructing relative clauses are possible in literary Tajik. In Tajik, however, there is also a broad range of nominalized and pre-nominal relative clauses that correspond partly to the Uzbek model, e. g. (Windfuhr and Perry 2009: 509):

- (98) *ob-e ki mo az hawz ovardem*
 water-*e* REL we from pool bring.PST.1PL
 ‘the water that we brought from the pool’
- (99) *ob-i mo az hawz ovardagī*
 water-*i* we from pool bring.PTCP
 ‘the water that we brought from the pool’
- (100) *az hawz ovardagī-mon ob*
 from pool bring.PTCP-1PL water
 ‘the water that we brought from the pool’

8.3. Temporal clauses

Persian distinguishes anterior, concomitant and future temporal clauses. One interesting feature is that anterior temporal clauses always employ the present subjunctive, even if the main clause is in the past:

- (101) *qabl az ānke ū be-ravad, man vāred*
 before from COMP he SBJV-go.3SG I entering
 šodam
 become.PST.1SG
 ‘before he went (away), I came in’

Future temporal clauses, on the other hand, usually require a past verb (but see §8.4.), even if the action lies in the future, e. g.:

- (102) *ba'd az ānke raftī Ālmān, mī-bīnī*
 after from COMP go.PST.2SG Germany DUR.see.2SG
ke ...
 that
 'After you (will) go to Germany, you (will) see that ...'

Concomitant temporal clauses are often introduced by the conjunction *vaqtīke* 'when':

- (103) *vaqtīke Aḥmad ānjā raft, dīd kasī nīst*
 when Aḥmad there go.PST.3SG see.PST.3SG nobody
nīst
 NEG.COP.3SG
 'When Aḥmad went there, he saw (that) there was nobody'

Vaqtīke may be replaced by *ke* moving forward through the sentence:

- Aḥmad ke ānjā raft, dīd kasī nīst*
Aḥmad ānjā ke raft, dīd kasī nīst
Aḥmad ānjā raft ke dīd kasī nīst

The first two versions express practically the same meaning as the sentence above with *vaqtīke*. In the 3rd version, with *ke* following the verb, the general meaning of the whole sentence remains similar, but a "subordinator switch" (Windfuhr and Perry 2009: 522) takes place and the second part of the sentence is syntactically focused: 'Aḥmad went there and saw: there is nobody'.

8.4. Conditional clauses

8.4.1. Impossible conditionals

There is a basic distinction between impossible and possible conditions in SNP. Impossible conditions in the present, or ones that have not been realized in the past, are expressed with the imperfect in the protasis and apodosis:

- (104) *agar mī-y-āmadam, mī-dīdam-aš*
 if DUR-GLIDE-come.PST.1SG DUR-see.PST.1SG-him
 'If I had come, I would have seen him', or: 'If I came (but I won't!), I would see him'

Past unrealized conditions can also be expressed with the pluperfect in the protasis:

- (105) *agar āmade būdam, mī-dīdam-aš*
 if come.PTCP COP.PST.1SG DUR-see.PST.1SG-him
 ‘If I had come, I would have seen him’

8.4.2. Possible conditionals

Possible conditionals fall into the following three categories:

- a. Conditions whose fulfilment is certain, or has already happened, employing the present indicative in the protasis:

- (106) *agar mī-ravī bīrūn, ū=rā mī-bīnī*
 if DUR-go.PRS.2SG outside he=rā DUR-see.PRS.2SG
 ‘If you go outside (I know you will!), you (will) see him’

- b. Conditions whose fulfilment is not certain, with the present subjunctive:

- (107) *agar be-ravī bīrūn, ū=rā mī-bīnī*
 if SBJV-go.PRS.2SG outside he=rā DUR-see.PRS.2SG
 ‘if / in case you go outside, you (would) see him’

- c. Conditions whose fulfilment is not certain, with the simple past; this emphasizes the immediate realization of the apodosis if the condition is fulfilled:

- (108) *agar raftī bīrūn, ū=rā mī-bīnī*
 if go.PST.2SG outside he=rā DUR-see.PRS.2SG
 ‘If you go outside, (right after doing that) you (will) see him’

8.4.3. Conditions without conjunction

In colloquial NP, any kind of condition may lack a conjunction, e. g.:

- (109) *mī-rī bīrūn, ū=rā mī-bīnī*
 DUR-go.PRS.2SG outside he=rā DUR-see.PRS.2SG
 ‘(If) you go outside (I know you will!), you (will) see him’

- (110) *raftī bīrūn, ū=rā mī-bīnī*
 go.PST.2SG outside he=rā DUR-see.PRS.2SG
 ‘(If) you go outside, (right after doing that) you (will) see him’

These sentences may alternatively be understood as concomitant temporal clauses ('when you go outside?'), with which they share an identical structure and sequence of tenses. The latter is actually the syntactically determining feature. Both types of subordinate clauses can be distinguished only by using the conjunction *agar* “if” or (*vaqtī*)*ke* “when”.

The close affinity of conditional and temporal clauses is also shown by the fact that in future temporal clauses with an indicative present in the main clause,

simple past and present subjunctive are equally employed in the protasis, the subjunctive introducing a doubt/condition into the temporal clause, e. g.:

- (111) *ba'd az ānke rafī / be-ravī Ālmān,*
 after from COMP go.PST.2SG / SBJV-go.PRS.2SG Germany
mī-bīnī ke...
 DUR.see.2SG that
 ‘After you (will / were to) go to Germany, you (will) see that ...’

8.5. Subordination vs. coordination

In modern colloquial Fārsī, many types of subordinate clauses can be constructed without the SNP multi-functional conjunction *ke* “that, which (etc.)”. Besides temporal, conditional (§8.3., §8.4.) and noun clauses (§8.1.), this concerns also relative clauses and clauses of purpose, e. g.:

- (112) *īn čī būd, xordī?*
 this what be.PST.3SG eat.PST.2SG
 ‘What was it (that) you ate?’
- (113) *ketāb-o xūndī, xoš-et ūmad?*
 book=rā read.PST.2SG well-your come.PST.3SG
 ‘The book (that) you read, did you like (it)?’
- (114) *Īrān rafte būdī čī kār konī?*
 Iran gone be.PST.2SG what work do.SBJV.2SG
 ‘You had gone to Iran, (so as) to do what?’

and other constructions, like the following, which may be understood concessively:

- (115) *xeilī zūr zad, na-šod*
 very power strike.PST.3SG NEG-become.PST.3SG
 ‘(Although) he tried very hard, it did not work’, or:
 ‘He tried very hard, (yet) it did not work’

The missing *ke* in these clauses turns a construction that is subordinate in SNP, into a paratactic sequence of clauses whose functional interdependency can only be understood from the context, and from the sequence of tenses. What seems to be a process of loss of *ke* as a clause conjunction in modern colloquial Fārsī, goes hand in hand with a de-grammaticalization of *ke* where it is still used. A widespread function of *ke* in colloquial NP is to emphasize the preceding word, without having any grammatical function, e. g.:

- (116) *man ke ne-mī-dānestam!*
 I EMPHAT NEG-DUR-know.PST.1SG
 ‘I really/actually didn’t know!’

This emphatic meaning of *ke* can be understood to replace its (former) grammatical meaning as conjunction in certain cases (see §8.3.2.), e. g.:

- (117) *man ke ānjā raftam, dīdam ...*
 I [that] there go.PST.1SG see.PST.1SG
 ‘when I went there, I saw ...’ > (preferred understanding:) ‘I actually went there, (and) saw ...’

See Mahmūdī Baxtiyārī et al (2013) for recent discussion of these additional functions of *ke* in contemporary spoken Persian. It remains to be seen if this “parataxization” of colloquial Fārsī syntax will affect also the grammar of SNP in the future.

Abbreviations

1, 2, 3	First, second, third person	IRR	Irrealis
COMP	Complementizer	NEG	Negation
COP	Copula	NUM	Numerical classifier
DEM	Demonstrative	OBJ	Object
DUR	Durative	PL	Plural
EMPHAT	Emphatic	PRS	Present tense
EZ	Ezāfe	PST	Past tense
GLIDE	Glide	PTCP	Participle
IMPER	Imperative	REFL	Reflexive
INDEF	Indefinite	SBJV	Subjunctive
INF	Infinitive	SG	Singular

References

- Ahadi, Shahram. 2001. *Verbergänzungen und zusammengesetzte Verben im Persischen: Eine valenztheoretische Analyse*. Wiesbaden: Reichert.
- Alam, Muzaffar. 1998. The pursuit of Persian: Language in Mughal politics. *Modern Asian Studies* 32(2). 317–349.
- Bani-Shoraka, Helena. 2005. *Language choice and code-switching in the Azerbaijani community in Tehran: A conversation analytic approach to bilingual practices*. Uppsala: Uppsala Universitet dissertation.
- Beeman, William O. 1986. *Language, status, and power in Iran*. Bloomington: Indiana University Press.
- Bonakdarian, Mansour. 2006. India. VIII. Political and cultural relations: Qajar period, the 19th century. *Encyclopædia Iranica* 8.1, 26–32. New York: Center for Iranian Studies, Columbia University.
- Borjian, Habib. 1388 [2009]. *Motūn-e Tabarī* [Tabarī texts]. Tehrān: Mīrāṣ-e Maktūb. (Āyene-ye Mīrāṣ, supplement nr. 15).

- Cole, Juan R. 2002. Iranian culture and South Asia, 1500–1900. In Nikki R. Keddie & Rudi Matthee (eds.), *Iran and the surrounding world: Interactions in culture and cultural politics*, 15–35. Seattle & London: University of Washington Press.
- Dabir Moghaddam, Mohammad. 1992. On the (in)dependence of syntax and pragmatics: Evidence from the postposition *-rā* in Persian. In Dieter Stein (ed.), *Cooperating with written texts: The pragmatics and comprehension of written texts*, 549–573. Berlin & Boston: De Gruyter Mouton.
- Durkin-Meisterernst, Desmond. 2014. *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Wien: Österreichische Akademie der Wissenschaften.
- Faghiri, Pegah, Pollet Samvelian & Barbara Hemforth. 2014. Accessibility and word order: The case of ditransitive constructions in Persian. In Stefan Müller (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo, 27–29 August 2014*, 217–237. Stanford, CA: CSLI Publications.
- Fragner, Bert. 1999. *Die "Persophonie": Regionalität, Identität und Sprachkontakt in der Geschichte Asiens*. Berlin: Das Arabische Buch.
- Gippert, Jost. 2007/08. Zur dialektalen Stellung des Zazaki. *Die Sprache* 47. 77–107.
- Gyselen, Rika. 2012. Scellements dans l'archive post-sassanide du Tabaristan (1). In Rika Gyselen (ed.), *Objets et documents inscrits en pârsig*, 97–108. Bures-sur-Yvette: Groupe pour l'Étude de la Civilisation du Moyen-Orient.
- Jazayery, Mohammad Ali. 2000. Farhangestān. In Ehsan Yarshater (ed.) *Encyclopædia Iranica* 9, 273–279. New York: Center for Iranian Studies, Columbia University.
- Jensen, Hans. 1931. *Neopersische Grammatik*. Heidelberg: C. Winter.
- Kahnemuyipour, Arsalan. 2003. Syntactic categories and Persian stress. *Natural Language & Linguistic Theory* 21. 333–379.
- Kauz, Ralph. 2009. Chang Zimei und der Islam in Shandong im 17. Jahrhundert: Akkulturation oder Abgrenzung? Zur Geschichte einer persischen Grammatik. *Saeculum* 60(1). 91–113.
- Khānlarī, Parvīz Nātel. 1979. *A history of the Persian language*, vols. 1, 2. New Delhi: Sterling Publishers et al.
- Khānlarī, Parvīz Nātel. 1365 [1986]. *Tārīx-e zabān-e Fārsī* [A history of Persian], 3 vols, (revised ed.). Tehrān: Našr-e nou.
- Kreiser, Klaus. 1993. Persisch als Schulsprache bei den osmanischen Türken: Von der Tanzimat-Zeit zur frühen Republik. In Jens Peter Laut & Klaus Röhrborn (eds.), *Sprach- und Kulturkontakte der türkischen Völker: Materialien der zweiten Deutschen Turkologen-Konferenz, Rauschholzhausen, 13.–16. Juli 1990*, 123–131. Wiesbaden: Harrassowitz.
- Landau, Jacob M. & Barbara Kellner-Heinkele. 2012. *Language politics in contemporary Central Asia: National and ethnic identity and the Soviet legacy*. London: Tauris.
- Lazard, Gilbert. 1956. Caractères distinctifs de la langue tadjik. *Bulletin de la Société Linguistique* 52(1). 117–186.
- Lazard, Gilbert. 1957. *Grammaire du persan contemporain*. Paris: Librairie C. Klincksieck. [Nouvelle édition: Téhéran: Institut Français de Recherche en Iran, 2006]
- Lazard, Gilbert. 1963. *La langue des plus anciens monuments de la prose persane*. Paris: Librairie C. Klincksieck.
- Lazard, Gilbert. 1966. L'enclitique nominal *-i* en persan: un ou deux morphèmes? *Bulletin de la Société Linguistique* 61(1). 249–264.

- Lazard, Gilbert. 1970. Étude quantitative de l'évolution d'un morphème: la postposition *rā* en persan. In David Cohen (ed.), *Mélanges Marcel Cohen*, 381–388. Berlin & Boston: De Gruyter Mouton.
- Lazard, Gilbert. 1971. Pahlavi, Parsi, Dari: Les langues de l'Iran d'après Ibn al-Muqaffa'. In Clifford Edmund Bosworth (ed.), *Iran and Islam, in memory of the late Vladimir Minorsky*, 361–391. Edinburgh: Edinburgh University Press.
- Lazard, Gilbert. 1982. Le morphème *rā* en Persan et les relations actancielles. *Bulletin de la Société Linguistique* 77(1). 177–208.
- Lazard, Gilbert. 1989. Le persan. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 263–293. Wiesbaden: Reichert.
- Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Costa Mesa: Mazda Publishers.
- Lehmann, Christian. 1978. *Yā-ye ešārat*. Zur Grammatik des persischen Relativsatzes. *Indogermanische Forschungen* 82 (1977). 97–106.
- Lenepveu-Hotz, Agnès. 2011. Sipāh-i Ispahān: devenir des groupes consonantiques initiaux moyen-perses en persan. *Faits de Langues* 38. 79–95.
- Lenepveu-Hotz, Agnès. 2014. *L'évolution du système verbal persan (X^e – XVI^e siècles)*. Leuven & Paris: Peeters.
- Liu, Yingsheng. 2010. A Lingua Franca along the Silk Road: Persian Language in China between the 14th and the 16th centuries. In Ralph Kauz (ed.), *Aspects of the Maritime Silk Road: From the Persian Gulf to the East China Sea*, 87–95. Wiesbaden: Harrassowitz.
- MacKenzie, David Neil. 1992. Codex Cumanicus. *Encyclopaedia Iranica* 5.8, 885–886. New York: Center for Iranian Studies, Columbia University.
- Mahmūdī Baxtiyārī, Behrūz & Farzāne Tāğ Ābādī. 2013. Čand kārbord-e kalāmī-ye dīgar az 'ke' dar Fārsīye goftār: Moṭāle'āt-e zabān va gūyešhā-ye ḍarb-e Īrān [Some other pragmatic functions of 'ke' in spoken Persian]. *Journal of Language and Western Iranian Dialects* 1(2), 1392 [2013]. 103–116. Kermanshah: Razi University.
- Marzolph, Ulrich. 2011. Das Schahname im Druck: Zu den lithographischen Ausgaben des persischen Nationalepos. In Julia Gonnella & Christoph Rauch (eds.), *Heroische Zeiten: Tausend Jahre persisches Buch der Könige*, 64–71. Berlin: SMB, Museum für Islamische Kunst, Staatliche Museen zu Berlin.
- Meunier, Annie & Pollet Samvelian. 1997. La postposition *rā* en persan: ses liens avec la détermination et sa fonction discursive. *Cahiers de Grammaire* 22. 187–232. Toulouse: Université de Toulouse – Le Mirail.
- Moinfar, Mohammad Djafar. 1980. Les classificateurs en persan. In G. Brettschneider (ed.), *Wege zur Universalienforschung*, 317–320. Tübingen: Narr.
- Nešāt, Mahmūd. 1368 [1990]. *Šomār va meqdār dar Fārsī* [Number and quantity in Persian]. Tehrān: Amīr Kabīr.
- Nawid, Senzil. 2012. Language policy in Afghanistan: Linguistic diversity and national unity. In Harold F. Schiffman (ed.), *Language policy and language conflict in Afghanistan and its neighbors: The changing politics of language choice*, 31–52. Leiden: Brill.
- Paul, Ludwig. 1998. The position of Zazaki among West Iranian languages. In Nicholas Sims-Williams (ed.), *Proceedings of the 3rd European Conference on Iranian Studies. Part 1, Old and Middle Iranian Studies*, 163–176. Wiesbaden: Reichert.

- Paul, Ludwig. 2002. A Linguist's Fresh View on "Classical Persian". In Maria Szuppe (ed.), *Iran. Questions et connaissances. Actes du IV^e Congrès Européen ...*, vol. 2, *Périodes médiévale et modern*, 21–34. Paris: Peeters.
- Paul, Ludwig. 2008. Some remarks on the Persian suffix =rā as a general and historical linguistic issue. In Simin Karimi, Vida Samian & Donald Stilo (eds.), *Aspects of Iranian Linguistics*, 329–337. Newcastle: Cambridge Scholars Publishing.
- Paul, Ludwig. 2013a. *A Grammar of Early Judaeo-Persian*. Wiesbaden: Reichert.
- Paul, Ludwig. 2013b. Persian Language i. Early New Persian. *Encyclopædia Iranica*. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/persian-language-1-early-new-persian> (accessed on 22 September 2017).
- Paul, Ludwig. (forthcoming). "Be nice to each other!" – Moral precepts at Tehran Metro walls as examples of written colloquial Persian. In Pavel Lurje (ed.), *Proceedings of the SIE conference held in St. Petersburg, 15–19 September 2015*.
- Perry, John. 1991. *Form and meaning in Persian vocabulary: The Arabic feminine ending*. Costa Mesa: Mazda Publishers.
- Perry, John. 2005. *A Tajik Persian Reference Grammar*. Leiden: Brill.
- Phillott, Douglas C. 1919. *Higher Persian Grammar*. Calcutta: University.
- Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian Phonological Systems*. Kraków: Nakł. Uniwersytetu Jagiellońskiego.
- Rāsex Mahand, Mohammad. 1386 [2007]. Touṣīf-e af’āl-e morakkab-e peybastī dar zabān-e Fārsī va ſīve-ye zabt-e ānhā dar farhanghā [The study of Persian enclitic compound verbs and their presentation in general dictionaries]. *Farhangnevīsī* 1(1). 236–252.
- Rypka, Jan. 1968. *History of Iranian Literature*. Dordrecht: D. Reidel Publisher.
- Sādeqī, ‘Alī Ašraf. 1363 [1984]. Tabdīl-e ‘ān’ va ‘ām’ be ‘ūn’ va ‘ūm’ dar Fārsī-ye goftārī [The replacement of 'ān' and 'ām' by 'ūn' and 'ūm' in spoken Persian]. *Majalle-ye Zabānšenāsī* 1(1). 52–71.
- Sādeqī, ‘Alī Ašraf. 1385 [2006–07]. Taḥavvol-e šāmet-e ‘q’-ye ‘Arabī dar zabān-e Fārsī [The development of the Arabic consonant 'q' in Persian]. *Majalle-ye Zabānšenāsī* 21(1–2). 3–32.
- Sādeqī, ‘Alī Ašraf. 1395 [2016]. Kasre-ye ezāfe be jā-ye ‘ke’-ye mouşūl [A short i instead of relative 'ke']. *Farhangnevīsī* 11. 3–7.
- Samvelian, Pollet. 2012. *Grammaire des prédicats complexes: Les constructions nom-verbe*. Paris: Hermès Science Publications.
- Shahidi, Minoo. 2008. *A Sociolinguistic Study of language shift in Mazandarani*. Uppsala: Uppsala Universitet dissertation.
- Spuler, Bertold. 1985. *Die Mongolen in Iran. Politik, Verwaltung und Kultur der Ilchanzeit 1220–1350*, 4. edn. Berlin: Akademie-Verlag.
- Tabātabāyī, ‘Alā’eddīn. 1395 [2016]. *Farhang-e touṣīfī-ye dastūr-e zabān-e Fārsī* [Descriptive dictionary of Persian grammar]. Tehrān: Farhang-e mo’āṣer.
- Tedesco, Paul. 1921. Dialektologie der westiranischen Turfantexte. *Monde Oriental* 15. 184–258.
- Windfuhr, Gernot Ludwig. 1979. *Persian Grammar: History and State of its Study*. Den Haag: Mouton Publishers.
- Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian Languages*, 416–544. London: Routledge.

4.7. Kumzari

Christina van der Wal Anonby

1. Introduction

Kumzari is a mixed language with Indo-Iranian and Arabian ancestors. The language is named after the village of Kumzar in Oman, which faces the Gulf from the Musandam Peninsula. Speakers of Kumzari live mainly in Kumzar and Khasab, with a few in scattered locations in Oman and the United Arab Emirates (see Figure 1). There is also a small population across the Gulf on Larak Island in Iran. In total, Kumzari speakers number approximately five thousand.

Kumzari's mixed nature places it within the Indo-Iranian, western Iranian family on the one side, and in the south Semitic family on the other side. An early brief description of the language was written by a British officer, Bertram Thomas, in 1929–1930. References to the people of Musandam speaking a different language appeared in earlier travelers' notes in the 19th century (Palgrave 1866; Ross 1874: 195; Jayakar 1902: 246; Zwemer 1902: 57; Miles 1994: 379, 436–448). From these accounts, and historical records of the region, as well as oral history of the people themselves (Lancaster and Lancaster 2011), a picture emerges of the origins of the Kumzari people.

The Musandam peninsula, in the ancient era called Maka or Mazun, was occupied intermittently by Persians of the Achaemenid, Parthian, and Sassanian dynasties, from the sixth century BC until the seventh century CE. During the Achaemenid rule of Maka, Aramaic was the lingua franca, and was used along with local languages to transmit messages from the Persian government (McWhorter 2007: 155; Potts 2012: 114). Several waves of Azd people fled Yemen after the breach of the Ma'rib dam and subsequent flooding from 145 BC and until the 5th century CE. Some Azd eventually settled in Musandam, with its capital at Dibba. The Azd migrants spoke a South Semitic language and settled among the Persian colonists in Oman.

There were wars and governing power went back and forth between the Azd and the Persians over seven centuries (Ross 1874: 116; al-Rawas 2000: 29; Ulrich 2008: 64). Dibba was a thriving centre of maritime trade in the third to seventh centuries CE, even called the capital of Oman, and Persian bishops were designated to represent Musandam to the Church of the East (de Perceval 1853: 14; Neale 1873: 132; Ross 1874: 75–79; Hawley 1970: 63; King 2001: 59–61). Within the Sassanian Empire, the Persians appointed Azd chiefs as their deputies to govern Oman. Northern Arabian tribes from Medina began raiding Oman in the 7th century. Some of the Azd chiefs joined them and war erupted. The Dibba wars eventually resulted in the expulsion of the ethnically mixed population in 633;

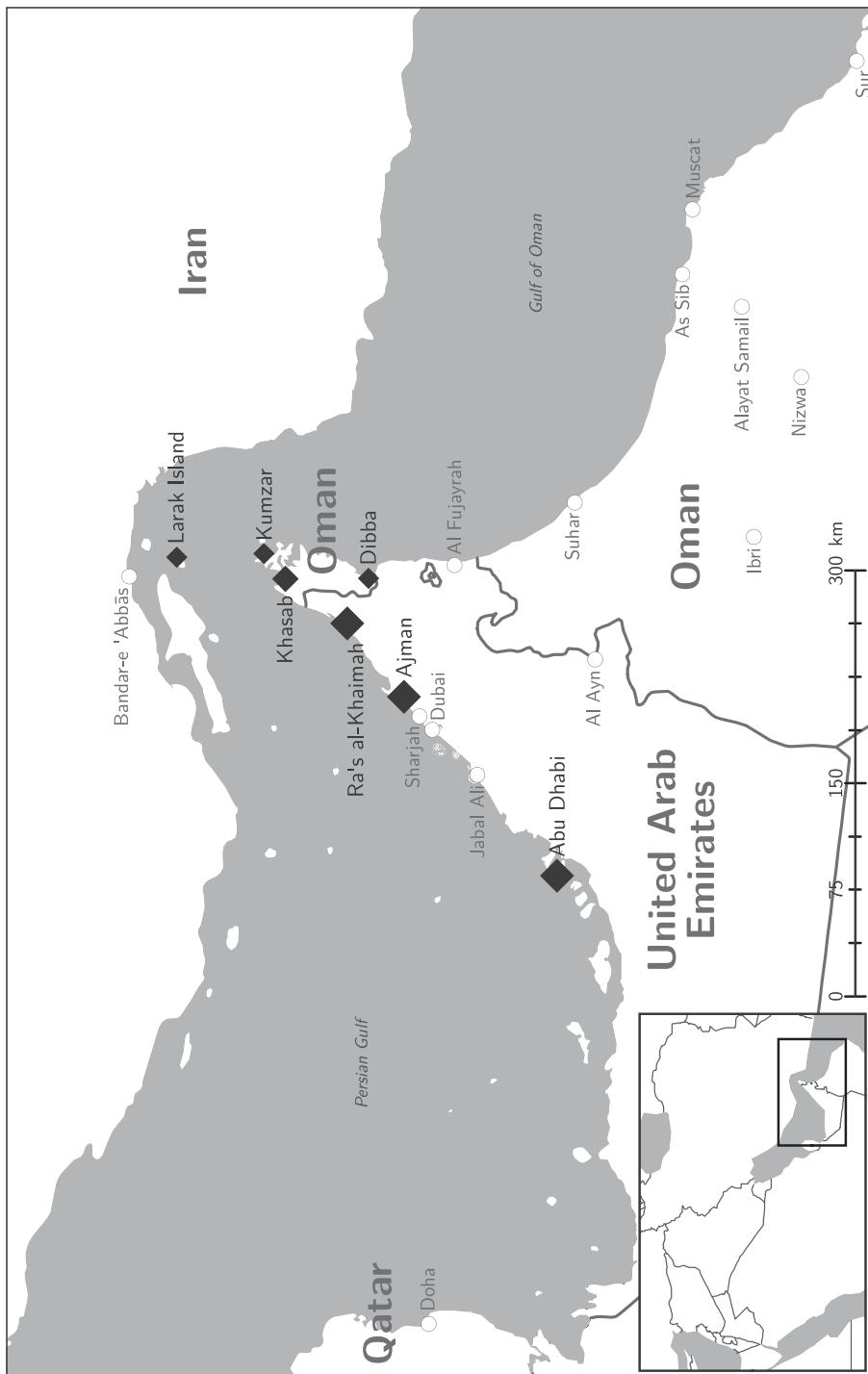


Figure 1: Locations of Kumzari speech communities

Persian and Azd rebels battling the North Arabians were defeated, killed, or chased from Dibba. The refugees settled in Musandam's mountains and remote coasts. Over the following centuries, they gained a reputation both as aloof bedouins and fierce guardians of the Strait between Arabia and Persia.

Linguistically, Kumzari resembles both the descendants of Middle Persian isolated from its modern counterpart in Iran, and the indigenous languages of southern Arabia. Core vocabulary comprises items of Semitic and Indo-Iranian sources, with both represented as well in the lexicon as a whole. Kumzari phonology maintains the complexity of Persian and Arabian inventories, including a full set of Semitic emphatic phonemes, even in lexical items of Iranian origin. The South Arabian retroflex approximant [ɻ], but also Iranian-specific sounds like the voiceless bilabial stop [p], the voiceless palato-alveolar affricate [tʃ], and the voiced uvular stop [g], are attested in the language. The verb system in Kumzari is fundamentally mixed, with both Arabian trilateral roots and Persian-origin finite verbs. Like South Arabian languages but unlike Iranian, Kumzari has post-constituent and multiply-marked negation. It has a split system of nouns, adjectives, and adverbs derived from Semitic verbs.

Kumzari has SOV constituent order similar to Iranian languages, but verb phrase syntax more closely aligned with South Arabian varieties. Within the Kumzari three-term system of evidentials, the firsthand sensory information source marker resembles particles with similar functions in Iranian Balochi dialects and South Arabian Mehri and Shihhi Arabic. Kumzari discourse incorporates features found in languages on east and west sides of the Gulf. Kumzari's mixed ancestry is reflected in the language as it is currently spoken.

Today the Kumzari consider themselves to be Arabs of the Shihuhu confederacy, with the identification of their unique language. Most Kumzaris lead a nomadic lifestyle. In the village of Kumzar, people subsist on fishing. They live in stone houses with central courtyards, with dwellings tightly wedged in the wadi, adjoined only by narrow paths. In summer months they travel by boat to the Kumzari quarter of the city of Khasab, where many have date palm plantations. They are there to harvest dates and trade goods and hold weddings. Kumzari are generally endogamous, but some intermarry with the Arabic-speaking Shihuh whose people make up the majority of the population on Musandam. The Kumzari are distinguished in lineage as part of a clan: Aqelī, Ghōšbanī, and Bō'īnī. By religion, the Kumzari follow the Sunni branch of Islam, and each of the two main clans has a mosque in Kumzar. Some religiously oriented oral traditions are adopted from Arabic, while Kumzari's own oral traditions include poetry, folktales, fishing and other work songs, lullabies, and *qawals*.

Kumzari is an unwritten language, but there is improvised use of the Arabic alphabet, such as to convey Kumzari text messages on mobile telephones. Kumzari local authorities approved an orthography developed by linguist Erik Anonby for use by Kumzari speakers in 2010, and there is some movement toward its

implementation in primary education. All but a few families teach their children Kumzari exclusively until school age. However, all media is in Arabic, and for schooling past primary grades children must move to Khasab for the full year. This presents a mixed prognosis for the future vitality of Kumzari: people are proud of speaking their own language, but it is in danger of being engulfed by Arabic in many domains of use. Most Kumzari speakers have some degree of fluency in Arabic, and those who remain in Khasab throughout the year are surrounded by speakers of the Shihhi dialect of Arabic.

Studies of Kumzari based on fieldwork are Jayakar (1902), Lorimer (1915), Thomas (1930), Dostal (1972), Najmabadi (1988), and Skjærvø (1989). The present author has written a Kumzari reference grammar, including discourse, texts, and lexicon (van der Wal Anonby 2015). Along with the author's collection, collaboration with a Kumzari speaker who has gathered several thousand words will form the basis of a dictionary.

2. Phonology

Kumzari has three short vowels /i/ /u/ /a/, which have a mid-centralised quality in comparison to its five long vowels /ī/ /ū/ /ē/ /ō/ /ā/. Vowels are always separated by a glide or glottal stop, even with clitics, and there are no vowel sequences. All vowel-initial words begin with a glottal stop.

Table 1: Vowel phonemes in Kumzari

	front	central	back
high	ī		ū
	I		U
mid	ē		ō
low		a	
		ā	

Kumzari distinguishes twenty-eight consonants. Its inventory draws both from its Persian heritage, in the phonemes /p/, /č/, and /g/, and from its Arabian heritage, as in its emphatic consonants: pharyngeal /ħ/, uvular /q/, and velarised alveolars /tˤ/, /dˤ/, /sˤ/, /zˤ/, and /lˤ/. Emphatics occur even in words not of Arabian origin, e. g. *sirx* 'red', *pānda* 'fifteen', *čāz* 'lunch'. All emphatic consonants in Kumzari have non-emphatic counterparts, with the notable exception of /zˤ/, which is always velarised.

Like Shihhi Arabic (Bernabela 2011: 26) and some South Arabian languages (Simeone-Senelle 1997: 381–382), Kumzari lacks certain phonemes from north-central peninsular Arabic varieties: the voiced pharyngeal fricative [ʕ] and

interdental fricatives [θ], [ð], and [d̪]. In common with the western Iranian languages Bakhtiari and southern Luri, Kumzari has a bilabial approximant /w/, in place of the Persian labiodental fricative /v/. Unlike Gulf Arabic and Persian but similar to the South Arabian languages Mehri and Hobyon and Shihhi Arabic, the Kumzari /r/ is a retroflex approximant; in some cases it can be realised as an alveolar or retroflex flap or trill (Bernabela 2011: 23–25; Simeone-Senelle 1997: 383).

Table 2: Consonant phonemes in Kumzari

	labial (-dental)	alveolar	velarised alveolar	(alveo-) palatal	velar	uvular	pharyn- geal	glottal
stop/affricate	p b	t d	t̪ d̪	č j	k g	q		'
fricative	f	s	š z	š		x ġ	ḥ	h
nasal	m	n						
approximant	w	l r	l̪	y				

Kumzari lacks sonority restrictions in its consonant clusters. Thus there are many words that violate the sonority sequencing principle (i. e. nasals, liquids, and approximants are found at the edges of a syllable, even outside a less sonorous consonant): *mṣaww* ‘barnacle’, *faql* ‘porcupine fish’, *rkāh* ‘sandal’, *ntōr* ‘treats’. This may be due to vowel elision in a reduced syllable, similar to the process in other Bedouin Arabic dialects that allow initial consonant clusters (Watson 2012: 901). All non-marginal consonants¹ are found in both first and second positions in initial and coda clusters. There are also no restrictions on word-internal consonant sequences across syllable boundaries. Most consonants in Kumzari form geminates, and consonant clusters separate at syllable boundaries.

The minimal syllable structure in Kumzari is CV. The following syllable shapes occur in monomorphemic words:

CV	<i>pi</i>	‘from’
CVV	<i>čō</i>	‘well’
CVC	<i>kaf</i>	‘palm, sole’
CVVC	<i>xōx</i>	‘peach’
CVCC	<i>bukr</i>	‘firstborn’
CVVCC	<i>qāpt̪</i>	‘white fish sp.’
CCVV	<i>xwā</i>	‘salt’
CCVVC	<i>qbīb</i>	‘narrow’
CCVCC	<i>mṣarr</i>	‘men’s headdress (turban)’
CCVVCC	<i>stārg</i>	‘star’

¹ I. e. all consonants except glottals, semi-vowels, and /!/.

Words of a single syllable and a short vowel must be closed; with a long vowel they may be open or closed syllables. Kumzari has penultimate syllable stress. Words with more than three syllables are rare.

Phonological processes that facilitate conformity to syllable limitations in Kumzari include assimilation, insertion, deletion, and resyllabification.

Word-final vowels *ū* and *ī* assimilate to *w* and *y* with the addition of a suffix: *gēdū* ‘water-pipe’, *gēdw-ē* ‘a water-pipe’; *qrādī* ‘bull shark’, *qrādy-ē* ‘a bull shark’. When the definite suffix *-ō* is added, *i* may be lowered to *ē* or backed to *u* and *a* raised and *u* lowered to *ō*: *battil* ‘dhow’, *battēl-ō* ‘the dhow’; *rikd* ‘corner’, *rukđ-ō* ‘the corner’; *langal* ‘anchor’, *langōl-ō* ‘the anchor’; *ğambur* ‘niche’, *ğambōr-ō* ‘the niche’.

Verb suffixes Realis *-d* and Perfect *-s* prompt the insertion of an epenthetic vowel *u* for verb roots ending in a single consonant and *ī* for verb roots with a final consonant cluster: *gnūn* ‘believe’, REAL *gnūnud*, PERF *gnūnus*; *gird* ‘go around’, REAL *girdid*, PERF *girdīs*. The same verb suffixes may affect the deletion of *r* in verb roots ending in that phoneme: *ēnar* ‘hide’, REAL *ēnid*, PERF *ēnis*, or the *r* may be retained as a flap (lenition) in the place of Realis *-d*: *ōdur* ‘hold on’, REAL *ōdur*, PERF *ōdus*.

Other cases of deletion in Kumzari are Imperfect prefix *t-* before alveolar-initial verb roots and the replacement of initial *w* with the Imperfect prefix: *jōr* ‘ask’, IMPERF *jōr*; *wāt* ‘want’, IMPERF *tāt*, as well as elision of initial *y* in an unstressed word following a short vowel-final word: *inda yē* ‘inside it’ > *indē*.

When a morpheme is added to a word ending with a consonant cluster, the final consonant joins the affixed syllable: *bukr* ‘firstborn’ CVCC, *buk.r-ō* ‘the firstborn’ CVC.CVV. Resyllabification also occurs with the Semitic prefix *mu-* on a syllable with a long vowel or a geminated coda; the *u* is deleted and the *m* becomes part of an initial consonant cluster in Kumzari: *mhāfid* ‘governor’, *mrād* ‘reason’.

Multiple phonological processes may occur to adhere to phonotactic constraints when suffixes are added, particularly in maintaining penultimate syllable stress. In the word *tarqit* ‘wedding poem’ with the definite suffix *tēruqt-ō* ‘the wedding poem’, *a* is raised and fronted to *ē*, *i* is deleted, and an epenthetic *u* is inserted.

3. Word Classes

Word classes include noun, pronoun, verb, deverb, adjective, demonstrative, quantifier, numeral, adverb, evidential, preposition, and clitics or particles functioning in negation, subordination, interrogative, existential, and discourse roles.

4. Nouns and Pronouns

Nouns in Kumzari have no case or gender, but are inflected with definite (*-ō*), indefinite (*-ē*), or plural (*-an*) suffixes, being generic by default: *qışr* ‘palace’, *qışr-ō* ‘the palace’, *qışr-ē* ‘a palace’, *qışr-an* ‘palaces’. Counted items take no plural suffix; instead the count suffix *-ta* is on the number: *gōsin* ‘goat’, *di-ta gōsin* ‘two goats’; its equivalent for counting humans is *kas*. The exception is inalienable nouns, which take both the count suffix *-ta* on the number and the plural suffix *-an*: *kōrk-* ‘son’, *af-ta kōrk-an* ‘seven sons’. Certain nouns can be abstract plurals, implying a state of affairs or a period of time, and similar semantics: *śartağ* ‘storm’, *śartağ-an* ‘season of stormy weather’; *aqil* ‘mountain’, *aqil-an* ‘mountainous region’.

Kumzari has a derivational suffix *-īn-* that turns a noun into an agent of that nominal property: ‘the one who (uses *n.*)’: *jāmal* ‘camel’, *jāmalīnō* ‘the camel-rider’. The resulting noun is obligatorily inflected with a definite, indefinite, or plural suffix. Another derivational suffix is *-ī*. It makes a noun into an adjective or abstraction of that nominal property: *indur* ‘inside’, *indurī* ‘inner’. A derivational suffix *-ītī* forms an adverb when appended to a noun, deverb, or adjective: *wuxr-ītī* ‘instantly’.

Pronouns follow a noun phrase to modify it, or can stand as independent pronouns. There are three personal pronouns, each with a singular and a plural form, indicated in the table below. The same forms are used as possessive pronouns following the noun: *xistar yē* ‘her fiancé’. Singular personal pronouns also have emphatic forms as shown in the following table.

Table 3: Personal pronouns

person	singular pronoun	plural pronoun	emphatic (S) pronoun
1	<i>mē</i>	<i>mā</i>	<i>ummē</i>
2	<i>tō</i>	<i>śmā</i>	<i>uttō</i>
3	<i>yē</i>	<i>śan</i>	<i>iyyē</i>

The reflexive pronoun *xō* is obligatory for an argument that agrees with the subject of the verb. Kumzari has a proximal demonstrative pronoun *yā*, a distal demonstrative pronoun *yē*, and a relative pronoun *ar*, as well as seven interrogative pronouns: *kē* ‘who’, *čē* ‘what’, *giya* ‘where’, *kāy* ‘when’, *čāb* ‘how’, *kārim* ‘which’, and *čanta* ‘how many’. In the absence of an interrogative pronoun, the clause-final interrogative enclitic *=ā* indicates a yes-no question.

Nouns can function as subjects, objects, or adverbial complements of movement verbs such as ‘go’, ‘come’, and ‘arrive’:

- (1) *gab twāra wābur ġuzr-ō.*
immediately shelter become:3S.REAL deep-the
‘He immediately took shelter in the deep [part of the well].’ (R1331)

5. Verbs and Deverbs

Kumzari has both verbs and deverbs, deriving from its dual heritage languages. Verbs are of Indo-European origin. Deverbs are of Semitic origin, usually in the form *CaCaCa*². Properties of deverbs are listed in the table below. Deverbs can form predicates with the existential enclitic or in a compound with a light verb.

Table 4: Word class properties of deverbs

Verb-like properties of deverbs	Non-verb-like properties of deverbs
<ul style="list-style-type: none"> – originate in Semitic verbs – require a subject and can take an object – underived forms cannot be subjects or objects – can be transitive or intransitive – when deverbs function as preverbal elements in compounds, deverbs take the object, thus following the same syntactic rule as a simple verb in Kumzari. – carry the semantic load of a verb in compounds – pronominal (existential) enclitics on deverbs might be understood to be equivalent to pronominal suffixes on verbs 	<ul style="list-style-type: none"> – do not take verbal inflections – are not etymologically related to Kumzari Indo-European verbs – do not have a complete conjugation into six TAMM forms like Indo-European verbs – require an existential enclitic or light verb to be a predicate – take the comparative suffix <i>-tar</i> – some deverbs express concepts that are typically associated with adjectives not verbs, e. g. <i>lakaša</i> ‘fat’, <i>rakka</i> ‘super’, <i>lawata</i> ‘weak’, <i>šaqšaqa</i> ‘funny’. – derive to form nouns, adjectives, adverbs

Compound verb constructions consist of a preverbal element followed by a light verb. The semantics of the compound rests in its preverbal element, which is most often a deverb drawn from an open class. In contrast, the light verb is one of two verbs distinguishing voice: *tka* ‘do’ for active and *tō'a* ‘become’ for passive, and these are inflected for aspect, mood, mirativity, person, and number.

5.1. Derived forms

Kumzari deverbs can be derived to form nouns, adjectives, and adverbs. Derived nouns and adjectives follow Semitic derivational morphology in which CV pattern

² Detailed information on Kumzari deverbs can be found in van der Wal Anonby 2015.

through stem alternation determines word class. Adverbs are derived from deverbs through the addition of a Kumzari suffix. Derived forms can be inflected according to their new word class.

Deverbs derive to nouns with deletion of the second vowel (or raising in quadrilateral roots) and addition of the suffix *-it* to the form *CaCCit*.

- baraza* (dv.) ‘appeared, appearing’
barżit (n.) ‘appearance’

Deverbs derive to adjectives in the form *CaCC*.

- qayama* (dv.) ‘stood, standing’
qaym (adj.) ‘upright’

Deverbs derive to adverbs by adding the suffix *-iti* (dropping the final *a* is morphophonemic).

- čaraxa* (dv.) ‘straddled, straddling’
čaraxītī (adv.) [e. g. sitting] ‘astride’

5.2. Verb forms

Kumzari has six verb forms: Realis, Perfect, Imperfect, Imperative, Irrealis, and Mirative. Aspect, modality, and mirativity are grammaticalised, while tense is coded lexically. Singular and plural of first, second, and third persons have separate morphemes.

Realis takes the form of a suffix *-d*. It is not a label of any temporal notion; rather, it is used for epistemic certainty, both for past completed events and certain future events.

The Perfect verb form, using the suffix *-s*, has a meaning of current relevance, especially resultative, of prior or complete events.

The Imperfect is a verbal prefix *t-* and labels ongoing or incomplete events. It is also used for progressive aspect, general statements, current statives, intended but unrealised future plans, the habitual, rhetorical questions, and purposive.

Imperatives are commands or requests. They are zero-marked, taking the pronominal suffixes directly on the verb stem, and the second person singular is zero-marked. Kumzari has a complete paradigm of Imperative forms for all persons and numbers.

The Irrealis also does not have distinct verbal inflection, but differs from other verb forms in its second person singular suffix *-ī*. The Irrealis verb form is used for the uncertain, such as future events which may not happen, and it covers semantic hypotheticality, potentiality, jussivity, conditionality, obligativity, and desiderativity.

The Mirative is a zero-marked morpheme (bare verb stem) for all persons. It

marks unexpected happenings or surprising information, including events with violence or magic.

Kumzari verb forms function on a scale of prominence in discourse, with foregrounding and backgrounding roles in structuring texts (see Hopper 1979). In narratives, the inciting incident and coda use Realis verb forms, the peak is punctuated with the Mirative, and the dénouement has both Realis and Mirative verb forms. Other discourse structures, including the introduction, nodus, and epilogue, generally consist of verbless clauses.

5.3. Verb inflection

Inflection affixes for verb forms, number, and person are shown the table below.

Table 5: Verb affixes

person	Realis	Perfect	Imperfect	Imperative	Irrealis	Mirative
1SG	-dum	-sum	t- -um	-um	-um	
2SG	-dī	-sī	t- -ī	Ø	-ī	
3SG	-diš/-d	-sē	t- -a	-a	-a	Ø
1PL	-dim	-sim	t- -im	-im	-im	
2PL	-dē	-sē	t- -ē	-ē	-ē	
3PL	-din	-sin	t- -in	-in	-in	

Verbs with simple morphology have one stem serving all verb forms. Verbs of more complex morphological type, namely *-ft* and *-št* verbs and *b-* and *w-* verbs, have one root to build the Realis and Perfect forms and a second simpler root as the basis of the Imperfect, Imperative, Irrealis, and Mirative forms.

Examples of simple and complex verb paradigms are given in the tables below.

Table 6: Example paradigm for the verb *fān* ‘send’

person	Realis	Perfect	Imperfect	Imperative	Irrealis	Mirative
1SG	<i>fāndum</i>	<i>fānsum</i>	<i>tānum</i>	<i>fānum</i>	<i>fānum</i>	
2SG	<i>fāndī</i>	<i>fānsī</i>	<i>tānī</i>	<i>fān</i>	<i>fānī</i>	
3SG	<i>fāndiš</i>	<i>fānsē</i>	<i>tāna</i>	<i>fāna</i>	<i>fāna</i>	<i>fān</i>
1PL	<i>fāndim</i>	<i>fānsim</i>	<i>tānim</i>	<i>fānim</i>	<i>fānim</i>	
2PL	<i>fāndē</i>	<i>fānsē</i>	<i>tānē</i>	<i>fānē</i>	<i>fānē</i>	
3PL	<i>fāndin</i>	<i>fānsin</i>	<i>tānin</i>	<i>fānin</i>	<i>fānin</i>	

Table 7: Example paradigm for the verb *xwaft*, *xwā* ‘sleep’

person	Realis	Perfect	Imperfect	Imperative	Irrealis	Mirative
1SG	<i>xwaftum</i>	<i>xwaftum</i>	<i>txwā'um</i>	<i>xwā'um</i>	<i>xwā'um</i>	
2SG	<i>xwaftī</i>	<i>xwaftī</i>	<i>txwā'ī</i>	<i>xwaw</i>	<i>xwā'ī</i>	
3SG	<i>xwaft</i>	<i>xwaftē</i>	<i>txwā'a</i>	<i>xwā'a</i>	<i>xwā'a</i>	<i>xwaw</i>
1PL	<i>xwaftim</i>	<i>xwaftim</i>	<i>txwā'im</i>	<i>xwā'im</i>	<i>xwā'im</i>	
2PL	<i>xwaftē</i>	<i>xwaftē</i>	<i>txwā'ē</i>	<i>xwā'ē</i>	<i>xwā'ē</i>	
3PL	<i>xwaftin</i>	<i>xwaftin</i>	<i>txwā'in</i>	<i>xwā'in</i>	<i>xwā'in</i>	

In addition to these, Kumzari has a large number of irregular finite verbs with idiosyncratic inflection paradigms. The most common irregular verbs are *ām* ‘come’; *dār*, *dō* ‘give’; *gid*, *ka* ‘do’; *gid*, *gir* ‘take’; *jīr*, *mēš* ‘see’; *raft*, *čō*, *rō* ‘go’; *wābur*, *tō* ‘become’.

5.4. Verb Phrase

In the verb phrase, verbs and deverbs follow the same syntactic rule: object complements in the form of a full noun precede the de/verb and those in the form of a pronoun follow the de/verb. In compounds, only a direct object in the form of a pronoun or an inalienable noun may come between the deverb and light verb.

- (2) Semitic deverb with noun object (object precedes deverb)

mīt waza'a tka ba diryī'tn-an.
fish distributing do:3S.IMPF to fisher-PL
'He distributes **fish** to the fishers.'

- (3) Indo-European verb with noun object (object precedes verb)

mīt dō'a ba diryī'tn-an.
fish give:3S.IMPF to fisher-PL
'He gives **fish** to the fishers.'

- (4) Semitic deverb with pronoun object (object follows deverb)

waza'a šan tka ba diryī'tn-an.
distributing 3P do:3S.IMPF to fisher-PL
'He distributes **them** to the fishers.'

- (5) Indo-European verb with pronoun object (object follows verb)

dō'a šan ba diryī'tn-an.
give:3S.IMPF 3P to fisher-PL
'He gives **them** to the fishers.'

Some Arabic varieties of Oman and the UAE have a similar distinction; that is, an object as a noun or a pronoun determines morphological marking on their verbal participles (Holes 1990: 48).

Each verbal complement is negated and subordinated separately in addition to the verb. Elements in a compound verb share negation and subordination.

Kumzari has several auxiliary verbs. Generally these occur with another verb to indicate modality. They precede the verb or compound, and the verb phrase syntax functions the same as a verb without an auxiliary.

Table 8: Auxiliary verbs

auxiliary verb	gloss
<i>wayda</i> ‘hold’	keep doing (one time)
<i>mād</i> ‘stay’	keep doing (habitual)
<i>čō</i> ‘go’	going to do
<i>tāta</i> ‘want’	want to or would do
<i>wēl</i> ‘let’	allow to do
<i>rāya, čwana</i> ‘be able’	can do

Verb goal arguments are clause-final and take no preposition. They encompass locative, benefactive, and instrumental complements. Factive verb phrases have the same syntax, with the entity that is “coming into existence” at the end of the clause. Factive syntax is used in the narrative exposition formula to introduce a character.

- (6) *raft mardk-ē wa zank-ē*
go:3s.REAL man-a and woman-a
‘There was a man and a woman.’ (U9)
- (7) *sīd-in šan madrēsit-ō.*
put:REAL-3P 3P school-the
‘They put them in the school.’ (B175)
- (8) *gardīd-iš xō tēr-ē.*
turn_into:REAL-3S REFL bird-a
‘He turned HIMSELF into a bird.’ (B231)

Kumzari has truncated verbal marking in subordinated clauses, serial verb constructions, and clause chains, where the full form marked with aspect, modality, mirativity, person, and number is in the main clause or final verb. Verbs in such multi-verb constructions generally share a subject.

6. Existentials

The existential enclitic is uninflected for tense, aspect, mood, voice, and mirativity properties, distinguishing only the person and number. It is clause-final and requires a complement in its function as a predicate in a clause. It is not compatible with another verb, suggesting that it may have emerged historically by dropping the **h*-verb stem, then broadening the pronominal endings' scope to accommodate all non-verbal predicates, at some point before Middle Persian adopted the *istad* forms. Existential enclitics are formally different from both pronominal verbal suffixes and possessive pronouns.

Table 9: Existential enclitics

person	singular	plural
1	= <i>um</i>	= <i>im</i>
2	= <i>ī</i>	= <i>ē</i>
3	= <i>Ø</i>	= <i>in</i>

Existentials cliticise to the end of any word class or phrase other than verbs; in doing so they undergo morphophonemic alternation, e. g. with an epenthetic glottal stop following a vowel.

Existentials can take existential, ascriptive, and identificational predicates (Pustet 2003). The three semantic types of predicates with an existential are given below.

- (9) [existential predicate]

ka pi yē si-ta=in ā, ka pi yē bātar!
if from 3s three-COUNT=EX:3P SUB if from 3s better
'If there were three of them, it would have been even better!' (K597)

- (10) [ascriptive predicate]

ammū šan zank-an=in ya’ni.
all 3P woman-PL=EX:3P that_is_to_say
'That is to say, all of them were women.' (S419)

- (11) [identificational predicate]

tō ahmad=ī? mē ahmad tka=um.
2s ahmad=EX:2s 1s ahmad "does-it"=EX:1s
'You are Ahmad?' 'I am Ahmad-Does-It.' (A110)

7. Negation

The particle of negation in Kumzari is *na*. The negative follows the constituent being negated. Post-constituent negation is not found in Iranian languages or Arabic. However, it is a “remarkable” distinguishing feature of the South Arabian languages of Oman (Waltisberg 2011: 319), and is possibly an innovation in these languages (Watson and Eades 2012: 3).

The negative is marked on the verb and on each post-verbal complement.

- (12) *mēy-ō fōšnīs-um ba šmā, jō’ar-ō fōšnīs-um na ba šmā na.*
 fish-the sell:PERF-1s to 2P pearl-the sell:PERF-1s NEG to
 2P NEG

‘I sold the fish to you; I did **not** sell the pearl to you.’ (K117)

Direct objects and deverbs in a compound are not negated separately.

- (13) *skafya k-ē na pi mē na.*
 concealing do:IMPER -2P NEG from 1s NEG
 ‘**Don’t conceal** from me.’ (S339)

An object in the form of a full noun precedes the verb, and thus does not take the negative. When the object is a pronoun, it follows the verb, and the negative particle then follows the object pronoun. The contrasting syntax between full noun and pronoun objects is shown in the following pair of examples:

- (14) *dar-ō twākš-um na.*
 door-the open:IMPF-1s NEG
 ‘I will **not** open **the door**.’ (S771)

- (15) *twākš-um yē na.*
 open:IMPF-1s 3s NEG
 ‘I will **not** open **it**.’ (S775)

The South Arabian language Mehri also varies negation syntax depending on whether a noun or pronoun is used (Rubin 2010: 265).

Prohibitive and negation of existential enclitics follow the same syntax: *na* follows the verb and each complement thereafter. In rejection, however, the negative particle precedes its referent.

- (16) *na bazzā wa na bīdar!*
 NEG beggar and NEG peasant
 ‘**Neither** beggar **nor** peasant!’ (B851)

8. Evidentials

Evidentials mark the source of information for a statement with grammatical, rather than lexical, meaning. They occur in clause-initial position, and are followed by the subordinating enclitic *=ā*, and then their complement in the form of a clause or noun phrase.

The three-term system of evidentials in Kumzari classifies the information source as firsthand sensory *tamna*, non-firsthand reportive *awa*, or non-firsthand inferred *ēka*. While non-firsthand sources are obligatorily marked, unmarked information is assumed to be firsthand. The extended function of *tamna* emphasises immediate observation or discourse peak.

The firsthand evidential emphasises information as being sensorily acquired, whether through visual, auditory, or tactile means, or even by emotions or premonition. In narrative texts, it can have presentative semantics or signal the discourse peak, and is often used to denote violent or magical happenings.

- (17) *tamna ā asp-ē inst̄ inda yē wa hawt-ē sirx.*
 SENS SUB horse-a humanlike in 3s and pool-a gold
 ‘He saw that a talking horse was in it and a pool of gold.’ (B351)

Reported information sources use the non-firsthand evidential *awa*, for both quotative and indirect speech.

- (18) *awa ā zīn-ō kišt-in.*
 REP SUB thief-the kill:REAL-3P
 ‘It is said they killed the thief!’ (R550)

Inferred sources of information are labelled with the non-firsthand evidential *ēka*. They include information obtained from deduction or general knowledge, and can convey pragmatic extensions of irony, deception, or disbelief.

- (19) *ar ġēla-an mā gis-ē ā, ēka ā yā grāb-ō.*
 that/who wheat-PL 1P take:PERF-3S SUB INF SUB this crow-the
 ‘The one who has taken our wheat, it must have been this crow.’ (G213)

Grammaticalised evidentials have not been reported for other languages of West Asia (as an extension of other categories see Johanson and Utas 2000), but are well-known in the Indo-Aryan family (Bashir 2006, 2010). The occurrence of potentially grammaticalised evidential particles has been suggested for languages of both Western Iranian and South Arabian families (van der Wal Anonby 2015). Evidentiality often diffuses in cases of intensive language contact (Aikhenvald 2004: 288).

9. Discourse particles

Discourse particles are clause-initial constituents that function grammatically to establish coherence in the structure of a text. The nine most commonly occurring plot structuring particles in Kumzari discourse are laid out in the following table. They are set on a scale of prominence that indicates their role in foregrounding or backgrounding information in a text.

Table 10: Discourse particles

more foregrounded (=high prominence)
<i>ka, amū, sā sā</i>
<i>byō</i>
<i>tamna</i>
<i>sā</i>
<i>čāb kin?</i>
<i>lumrād, filhāl</i>
more backgrounded (=low prominence)

At the discourse peak of a narrative, clusters of foregrounding particles occur.

- (20) *bārē gid-iš, dikkara sātē. sā sā sikkara ba yē,*
once do:REAL-3s twice now right.away thrice to 3s
wa
and
'He did it once, twice now. **Right away**, as he was going around the third time, and
- (21) *bast ya 'nī, wa ka dakka pā kin bā*
finishing that_is_to_say and quickly digging foot do:MIR on
yē
3s
he is just finishing, I mean, and **quickly** he [the boy] stands firmly
- (22) *zamy-ō. byō rāstāg yē ka byō dist*
ground-the come:MIR straight 3s suddenly come:MIR hand
xō
REFL
on the ground. **Immediately** he [the sorcerer] comes in line with him, **suddenly** he [the boy] comes and puts his hands

- (23) *sō zekon yē. yē wa ka kēsaft kin*
 put:MIR backside 3s 3s and suddenly plunging do:MIR
inda yē,
 inside 3s
 on his [the sorcerer's] backside. **Suddenly** he comes to him, and **suddenly** he plunges him into it,
- (24) *inda qiz'an-ō.*
 inside cauldron-the
 into the cauldron.' (B560)

10. Prepositions

Kumzari prepositions precede their complements. Prepositions are separate words in Kumzari, except *wā* which can cliticise to its complement.

Table 11: Prepositions

preposition	gloss
<i>ba</i>	to, for, of, on, with
<i>bağā</i>	without
<i>bar</i>	belonging to
<i>ğay</i>	except
<i>hata</i>	until
<i>inda</i>	in, inside
<i>mayya</i>	between, among
<i>naxa</i>	aboard (a vehicle)
<i>pana</i>	beside, near
<i>pi</i>	from, than, since
<i>sīna</i>	toward
<i>tē</i>	before
<i>wā</i>	with, at, -ward

10.1. Proclitic directional and possessive preposition *wā*

The preposition *wā* is a directional, with a locational noun as its complement. In a directional function, *wā* cliticises to its complement.

- (25) *č-in wā=ēbar wa tā'-in*
 go:IMPF-3P at/-ward=over_here and come:IMPF-3P
wā=ēbar.
 at/-ward=over _here
 'They would go **here** and **there**.' (R309)

Kumzari also uses the preposition *wā* with a complement as a possessive construction, in place of a ‘to have’ verb. The complement is the possessor in the form of a noun or pronoun, and follows the preposition. A similar possessive construction occurs in the South Arabian languages, where it is *d-/d-* with a pronoun complement (Simeone-Senelle 1997: 386). In the Kumzari possessive construction, *wā* does not lose stress to its complement as it does when it functions as a directional. The possessive *wā* does not have verbal inflection for categories such as aspect, mood, or mirativity. The possessive construction falls at the end of a clause, unless it is a factive (Payne 1997: 59).

- (26) *sā řēx-ō ā, jāmal wā yē.*
now sheikh-the TOP camel with 3s
'Now, the sheikh, he **had** a camel.' (A397)

Like factive verb syntax, when it denotes the coming into existence of an entity, the *wā* construction occurs clause-initially.

- (27) *raft řēx wālēyit-ō, wā yē si-ta ditk-an.*
go:3s.IMPF sheikh country-the with 3s three-COUNT daughter-PL
'There was a sheikh of the country; he **had** three daughters.' (S12)

11. Adjectives

Adjectives in Kumzari can function as attributes or predicates. However, attributes are more often expressed by nouns, deverbs, abstract plurals, and verbs in Kumzari.

In a noun phrase, adjectives occur after the noun they modify and before its possessive noun or pronoun. In this role, they agree in definiteness and number with the noun they modify.

- (28) *qēšarit-ō jwān-ō ġaraqa gid-iš.*
date-the good-the swallowing do:REAL-3s
'He swallowed **the** good date.' (N98)

As predicates, adjectives occur with a following existential enclitic or verb *tō'a* ‘become.’

- (29) *čēnağ=in.*
thirsty=EX:3P
'They are **thirsty**.' (R1220)

Adjectives can take nominal suffixes such as definite *-ō*, indefinite *-ē*, and plural *-an*; so inflected they can function syntactically as a noun in a clause argument:

- (30) *gap-an* *raft-in* *xāna* *ba* *rōr-an* *āmō* *xō.*
 big-PL go:REAL-3P marriage with child-PL uncle(paternal) REFL
 ‘The older ones [literally: ‘bigs’] married their cousins.’ (S27)

Adjectives may be derived from other constituents, such as nouns, with the suffix *-ī*.

<i>awwēlī</i>	‘first’	<i>awwal</i>	‘first, firstly’
<i>wardī</i>	‘pink’	<i>ward</i>	‘flower’
<i>zēranī</i>	‘lower’	<i>zēran</i>	‘bottom’

Other adjectives are derived from deverbs with the Semitic triliteral root following the pattern CaCC:

<i>ğayb</i>	‘absent’	<i>ğayaba</i>	‘finishing with’
<i>haml</i>	‘cautious’	<i>hamala</i>	‘heeding’
<i>tafṣ</i>	‘rotten’	<i>tafasa</i>	‘rotting’

The triliteral stem of the adjective is *CaCaCa* when it takes the comparative suffix *-tar* (lengthening of the final vowel in the stem is morphophonemic).

<i>balağā</i>	dv.	‘ripening’
<i>balğ</i>	adj.	‘ripe’
<i>balağātar</i>	adj.	‘riper’

12. Demonstratives

Kumzari has two demonstratives: proximal *yā* and distal *yē*. Unlike possessive pronouns, they precede the noun they modify, and must accompany it; they cannot stand alone. The noun they modify takes the definite suffix *-ō*.

- (31) *yē* *čāb* *kin*, *yē* *şāħar-ō* *ā?* *gardīdiš*
 3s how do:MIR DEM sorcerer-the INTERR turn_into:3s.REAL
xō *tēr-ē*.
 REFL bird-a
 ‘How did he do it, **that sorcerer**? He turned himself into a bird.’ (B228)

13. Quantifiers

Quantifiers precede the noun they modify, expressing an amount.

- (32) *nikta arma dī'-in ba mā.*
 a.bit date give:IMPF-3P to 1P
 ‘They would give **a bit of dates** to us.’ (K48)

Common quantifiers are listed in the table below.

Table 12: Quantifiers

<i>ay</i>	‘any’
<i>iš</i>	‘none, any’
<i>dasta</i>	‘a number’
<i>nikta</i>	‘a bit’
<i>kam</i>	‘a little’
<i>habba</i>	‘a few’
<i>qadar</i>	‘some’
<i>ammū</i>	‘all’

14. Adverbs

Kumzari has two types of adverbs with differing morphosyntax: clause-initial adverbs modify the entire clause, and post-verbal adverbs relate only to the verb.

- (33) [clause-initial adverb]
quṭ gōsin mē jīs-ī ā?
 ever goat 1s see:PERF-2S INTERR
 ‘Have you **ever** seen my goat?’ (N28)

- (34) [post-verbal adverb]
tiyar tī'-im šarg.
 finished become:IMPF-1P quickly
 ‘We’re **finishing quickly**.’ (U559)

Post-verbal adverbs can only be negated when the verb they modify is negated, but clause-initial adverbs can be separately negated. Like other adverbs, derived adverbs with the suffix *-ītī* follow the verb they modify.

15. Clauses

A minimal clause in Kumzari is a subject and a predicate. However, the subject may be implicit, and the predicate may be a verb phrase, an existential enclitic, an evidential, a possessive construction, or an abstract plural noun. Subject-Object-Verb is the unmarked constituent order, but topicalisation or factivity may reposition an argument.

- (35) *āmō mē tālbū tō gis-ē ba ēnar-ō.*
 father-in-law 1s request 2s do:PERF-3S for henna-the
 ‘My father-in-law has requested you for [doing] the henna.’ (S436)

In clause-initial position there may be discourse particles, evidentials, or adverbs pertaining to the entire clause. In clause-final position there may be prepositional phrases and verb goal arguments.

- (36) *sā sōr wa pīma tk-um ba šmā čāz.*
 now salt_fish and green_onion do:IMPF-1S for 2P lunch
 ‘Now I will make salt fish and green onions **for you for lunch.**’ (G865)

Nominal, adjectival, and locational predicates require an existential enclitic.

- (37) *šēx=īn na ā, ādī=īn.*
 sheikh=EX:3P NEG SUB normal=EX:3P
 ‘**They were** not royalty; **they were** common.’ (U12)

- (38) *halla ba tā'r=īn.*
 landing on mountain-ledges=EX:3P
 ‘**They are** landing on mountain-ledges.’ (B599)

Existential enclitics attach to predicates consisting of perfect participles, or deverbs not in a compound.

- (39) *āwwa bār maxlōq-an ġafala=īn.*
 first time people-PL unsuspecting=EX:3P
 ‘The first time the people **were** unsuspecting.’ (A166)

- (40) *bukrit-an zās-in=īn.*
 kid-PL give_birth:PERF-3P=EX:3P
 ‘The kids **are** born.’ (N24)

15.1. Complex clauses

15.1.1. Coordination

Coordinated clauses in Kumzari are linked by conjunction, disjunction, adversative coordination, or causal coordination. Each has its own set of linking forms.

Conjunction uses *wa* ‘and’ or *ka* ‘also’ between clauses, with their subjects having coreferential or disjoint reference.

- (41) *hubbō-ō* *xaṭt-ē* *kataba* *gid-iš*.
 grandmother-the message-a writing do:REAL-3s
 ‘The grandmother wrote a message.’
- (42) *wāqā* *gid-in* ***wa*** *fānd-in* *yē*.
 signature do:REAL-3P and send:REAL-3P 3s
 They signed **and** they sent it.’ (P348)

The consecutive conjunction *ka* preceding each phrase marks listing parallelism with multiple coordinands (Haspelmath 2007: 15).

- (43) *šaw* *ā*, ***ka*** *maylat* *gid-in*, *wa* ***ka*** *ammū*
 night SUB LIST wedding.poetry do:REAL-3P and LIST all
 ‘At night, **also** they did the Maylad [poetry], and **also**
 čī *gid-in* *wa* ***ka*** *srō* *kēšid-in*
 thing do:REAL-3P and LIST sung_poetry pull:REAL-3P
 they did everything, and **also** they sang the Sro [poetry],
 wa *tiya* *būr-in*.
 and finished become:REAL-3P
 and they finished.’ (S632)

Causal coordination uses *ka* alone.

- (44) *xalaqa* *ba* *yē* ***ka*** *tāt-um* *dig-um* *yē*
 (good)_looks to 3s so want:IMPF-1s take:IMPF-1s 3s
 šū-ī.
 husband-ADVR
 ‘He is handsome, **so** I want to take him as a husband.’ (B773)

Disjunction uses *wala* ‘or’ between the two elements, or the emphatic form *wana* ‘either, or’ preceding each clause for multiple coordinands (Haspelmath 2007: 15).

- (45) *kam* *ḡāz* *dō-um* *ba* *yē* ***wala*** *śirx* *tāt-a*
 how_much money give:IMPF-1s to 3s or gold want:IRR-3s
 ā,
 SUB
 ‘I will give her however much money, **or** if she wants gold,

- (46) *sirx dō-um ba yē.*
 gold give:IMPF-1s to 3s
 I will give her gold.' (P468)

Two other counterfactual conditionals used in Kumzari are *kana* and *midam*, both meaning ‘otherwise’.

Kumzari has the adversative conjunction *lakin* ‘but’, and the more rarely heard *balkē* ‘however’.

- (47) *ka byō̄ ba yē lakin adliy-ē ba yē.*
 right_away come:MIR to 3s but gown-a to 3s
 ‘Right away he came to her **but** she was wearing a gown!’ (S426)

Clauses can also be linked asyndetically, with only juxtaposition and rising-falling intonation to indicate coordination.

15.2. Subordination

Subordination in Kumzari uses four types of multiple verb constructions: relative clauses, clause chains, adverbial clauses, and complement clauses. All signal clause relationships through the subordinating enclitic =ā, which occurs clause-finally and has allomorphs =ō̄, =wā̄, and =yā̄.

- (48) *bard-in yā̄ jitt-ō̄ ā̄, dakka yē gid-in*
 carry:REAL-3P this corps -the SUB burying 3s do:REAL-3P
inda maqbart-ō̄.
 in grave-the
 ‘Carrying this corpse, they buried it in the grave.’ (A290)

In a subordinated clause, the subordinator is obligatorily marked on every complement of the verb. The enclitic ā follows both the verb and the verbal complement.

- (49) *wa āmad-in ā̄ ba xwā̄ ā̄, bang-an.*
 if/when come:REAL-3P SUB for salt SUB dusk-PL
 ‘When they came for the salt, it was nightfall.’ (B649)

Adverbial clauses and complement clauses take the subordinating enclitic ā clause-finally.

- (50) *tē ba rēs-a ba y'=ā̄, grab-ō̄ pōrid.*
 before to arrive:IRR-3S to 3S=SUB crow-the fly:3S.REAL
 ‘Before he reached it, the crow flew away.’ (G221)

- (51) *mār, aqrab inda yē ā̄, dām na.*
 snake scorpion in 3S SUB know:1S.IMPF NEG
 ‘I don’t know [whether] there was snake or scorpion in it.’ (P110)

The subordinator follows all evidentials: sensory, reported, and inferred.

- (52) *sā wa barza wāb ba bāg almw̥z*
 now if/when appearing become:MIR to garden AR.: (the-banana)
 \bar{a} ,
 SUB
 ‘Now when he showed up at the banana garden,
 (53) **tamna** *ā yē haps=in.*
 SENS SUB 3s bound=EX:3P
he saw that his brothers were bound.’ (B1138)

The subordinating enclitic *ā* also joins semi-subordinate constructions such as medial clauses in clause chains and repeated clauses in anadiplosis, and it can follow a noun phrase as a topicaliser. Consecutive linked clauses in logical or temporal sequence use the rare subordinator *bēw* in place of *ā*.

Other clause subordination is achieved through the irrealis verb form, perfect participles, and deverbs, taking advantage of their atypical properties to form relative, complement, or adverbial clauses.

Subordinated clauses may shorten the verb stem by reducing inflection for aspect, mood, and mirativity. This strategy is used for adverbial and complement clauses, and for serial verb constructions such as chains, where only the final verb contains aspect, mood, and mirativity information.

15.2.1. Relative Clauses

For forming relative clauses, Kumzari has a clause-initial relative pronoun *ar* (post-nominal allomorph /a/) and a clause-final relative particle *na* (the latter can be fused to the subordinating enclitic as in the example below).

- (54) *ar jamma tk-a z̥īn-ō, jamma tk-a,*
 that/which/who bending do:IMPF-3S thief-the bending do:IMPF-3S
‘The one who bends down is the thief, [the one who] bends down,
sayya ba xō n=ā, z̥īn-ō.
 picking_up for REFL of_which=SUB thief-the
who picks up, is the thief.’ (A641)

Relative clauses in Kumzari may also use a gap strategy, with a nominalised verb in the form of a deverb or perfect participle.

- (55) *xāṭt-ē āma ba mē, law yē haraqā.*
 message-a come:3s to 1s edge 3s sealed_with_wax
 ‘A message came to me, [**of which**] its edge was **sealed with wax**.’
 (P1099)

Prepositional phrases with *pi* ‘from’ are commonly employed as a relativisation strategy, similarly to the endemic use of a relativiser *bū* (in place of *illi*) in Arabic dialects of northern Oman (see Eades 2009: 91).

- (56) *fān yē ba rōk-ō, pi raft-ē hijj.*
 send:2S.IMPER 3s to boy-the from go:PERF-3S hajj_pilgrimage
 ‘Send it to the boy, **who** has gone on the Hajj pilgrimage.’ (P530)

15.2.2. Clause Chains

Kumzari has clause chains that link several clauses of co-subordinated, truncated medial verbs each followed by the subordinating enclitic =ā; an independent final clause with a fully inflected verb takes the aspect-mood-mirativity marking for the entire chain. The clauses in a chain share the same subject and are linked semantically, such as in an action sequence. Serial verb constructions are known to occur in many languages of western Iran (Stilo 2004: 294, 296).

Anadiplosis in Kumzari repeats a preceding clause in a subordinated clause with the enclitic =ā, as presupposed information in a discourse.

- (57) *ka jahha kin ba rōk-ō awēlī, rōk-ō jwān-ō.*
 PEAK swooping do:MIR for boy-the first boy-the fine-the
 ‘Immediately **he swooped for the first boy**, the fine one.
- (58) *jahha kin ba yē ā wa sayy yē pi*
 swooping do:MIR for 3s SUB and lift_up:MIR 3s from
mayya rōk-an.
 midst boy-PL
He swooped on him and lifted him up from among the boys!’ (B238)

Kumzari narratives have embedded poems with FRAME and LIST structure (van der Wal Anonby 2009). The poems often begin with anadiplosis in the frame, with the list being verbless or using Irrealis or truncated verbs with the enclitic =ā, similar to clause chains.

15.2.3. Adverbial Clauses

Adverbials in Kumzari encompass time, location, purpose, reason, circumstantial, simultaneous, additive, and conditional clauses. Conditional clauses use the clause-initial subordinating conjunctions *wa* (reality), *ka* (unreality), *wana* (predictive), or *kana* (counterfactual), with varied meaning depending on the verb forms in the joint clauses and whether the subordinate clause contains the enclitic =ā.

- (59) *ēka ā y'=ā kas tāt-a yē na, kaft -ē*
 INF SUB DEM=SUB PERS want:IMPF-3S 3S NEG fall:PERF-3S
ba čāf-ō bē.
 on beach-the only
 ‘Obviously no one wanted this, [since] it was just left on the beach.’
 (S244)
- (60) *hata tay wā=bāla ā, gōsin-an txōr-in.*
 until come:3S.IMPF -ward=up SUB goat-PL eat:IMPF-3P
 ‘The goats were eating until he came up.’ (G142)
- (61) *āxur bang šmā bzēn-um ā, arzamē u yēkē*
 after calling 2P hit:IMPF-1S SUB one_at_a_time
byāt.
 come:3S.IRR
 ‘After I call out to you, [you] should come one at a time.’ (R455)

15.2.4. Complement Clauses

Kumzari allows for subject, object, and oblique complements. Complement clauses may take truncated verbs with reduced aspect-mood-mirativity marking, or a diminished form such as irrealis, deverbal, or perfect participle. Each complement of the verb is marked with the clause-final subordinating enclitic =ā, and negation on the matrix verb and the entire complement clause. A clause-initial complementiser used specifically for oaths is *inna*.

- (62) *şwāl mām xō gid-iš ka jilbē yē*
 question mother REFL do:REAL-3S if/when scarf 3s
čišt-ē.
 wash:PERF-3S
 ‘She asked her mother whether she had washed her scarf.’ (N20)
- (63) *ātiš-ē jīr-iš, ātiš-ē āntē čō-ō ā, ātiš-ē labaqa.*
 fire-a see:REAL-3S fire-a there well-the SUB fire-a burning
 ‘He saw a fire, a fire that was there by the well, a burning fire.’ (R360)
- (64) *dit mē tal bur tō gis-ī*
 daughter 1s decided become:3S.REAL 2s take:PERF-2s
xustār-ī.
 fiancé-ADVR
 ‘My daughter has decided that you be taken as a fiancé.’ (B931)

16. Discourse

Kumzari discourse uses particles, formulae, verb forms, and other means to structure a text and achieve coherence.

16.1. Formulae

There are two kinds of formula in Kumzari: narrative and thematic. Narrative formulae have conventional wording and occur obligatorily once in every narrative text. Thematic formulae recur throughout the text but are idiosyncratic for each narrative.

There are only four narrative formulae: two in the exposition (*qıssıte wa hakāy-itē* ‘a story and a telling’, *raft yēkē* ‘there went someone’), and two in the conclusion (*tō raftī wa mi āmadum* ‘you went and I came’, *xalaş* ‘the end’). Thematic formulae are closely linked to content in a narrative (Kossmann 2000: 75), and may resemble proverbs. A single thematic formula specific to a particular narrative is repeated at several points in the text.

16.2. Speech Type

Each section of the plot in narrative features quoted speech, drama speech (without quote margin e. g. ‘she said’), or no speech. Like verb forms, speech types function on a scale of prominence aligned with their grounding role in discourse. The exposition and conclusion of a tale do not have speech. In the body of the tale, information is foregrounded with drama speech, or backgrounded through quoted speech.

16.3. Plot structure

The narrative plot in Kumzari is divided into eleven narremes, each within one of three sections: the exposition, body, and conclusion. A summary of Kumzari narremes and their content is in the table below. Each narreme in plot structure is characterised by the use of certain discourse particles, verb forms, formulae, speech types, and other features.

Table 13: Summary of plot structure narremes

exposition:	
aperture	“a tale. there went someone.”
introduction	characters and setting
nodus	presentation of a problem
body:	
inciting incident	the first happening that turns the course of the story
intentus	development of tension, several pivots
accalmie	calm before the storm
peak	highest point of conflict
dénouement	unravelling of tension
conclusion:	
coda	resolution of central conflict, homecoming
finis	“you came, I went. the end.”
epilogue	“was it good? that’s how it happened.”

16.3.1. Exposition

The aperture contains one or both of the narrative formulae opening the tale. The introduction gives background information in the form of characters and setting. It is generally verbless, with much lexical repetition and paraphrase.

The nodus extends the setting before any pivot in the tale, by presenting the thematic problem, complication, or conflict; it usually focuses on wealth or lineage. The nodus is generally speechless and verbless and contains much repetition and parallelism.

16.3.2. Body

The inciting incident is the initial event that directs the course of the narrative. It begins immediately with Realis verb forms and the text’s first instance of direct speech and culminates in the main character leaving home (see “absentation” Propp 1968: 26). The intentus is a series of ascending pivots, with straightforward description of action and little repetition. Grounding develops tension in this narreme.

Pivots are initiated with the discourse particles *tamna* and *bīyō* and conclude with *lumrād*. The accalmie provides a lull in the action, containing backgrounded information that draws out tension to highlight the peak. There is much repetition, lists, embedded poems, and the low-prominence discourse particles *sā* and ‘čāb kin?.

The peak represents the high point of the plot, with the opening of conflict presented in the nodus. Speech is replaced by action and there is the “crowded stage”

effect with all characters present. The evidential *tamna* and Mirative verb forms are used along with foregrounding discourse particles *ka*, *amū*, and *sā sā*. After the peak, the dénouement cools the conflict. It still has Mirative and Realis verb forms, but features quoted speech prominently and uses backgrounding discourse particles *lumrād* and *filhal* to summarise information.

16.3.3. Conclusion

The coda resolves the nodus, with the central character returning home in success. It uses Realis verbs, no speech, and often includes an explanation directed to the audience. The coda is followed immediately by the finis formulae, and then an interactive epilogue that summarises the narrative theme (cf. Utas 2006: 227).

17. A text from the Kumzari folktale Sōntyō

- (65) *rōz-an dgur nwāšam y'=ā, ra ba yē dgō*
 day-PL next evening 3s=SUB go:3S.REAL to 3s say:3S.IMPF
ba yē
 to 3s
 ‘In the evening of the next day, he went to her, he said to her,
- (66) *maš, hā! wana dar-ō wākiš, wana nwāz*
 see:2S.IMPER well either/or door-the open:2S.IMPER or_else
 tomorrow
 ‘See here. Either you open the door, or else tomorrow
- (67) *nijjar-an tār-um, lōh-ō šaraxa tk-in.*
 boatbuilder-PL bring:IMPF-1S wood-the chopping do:IMPF-3P
taftafa yē tk-im
 smashing 3s do:IMPF-1P
 I will bring the boat builders, [and] they will chop the wood. We will smash it
- (68) *ba ēmag. šan wā maqta-an šan wa mīšar-an šan,*
 for firewood 3P with chisel-PL 3P and saw-PL 3P
wa ādamī
 and person
 [the boat] into firewood! They, they have their chisels and their saws. And people

- (69) *txēn-in bā tō. dar-ō wākid-iš ba yē.*
 laugh:IMPF-3P against 2s door-the open:REAL-3S for 3s
 will laugh at you.” She opened the door to him.
- (70) *dgō ba yē maš, hā! mē ūan tō=um.*
 say:3S.IMPF to 3s see:2S.IMPER well 1s wife 2S=EX:1s
gid-ī mē,
 take:REAL-2S 1s
 She said to him, “See here. As for me, I am your wife. You married me,
- (71) *wa lakin tāt-ī tēbur-ī mē wā= xā- šmā*
 and except want:IMPF-2S carry:IMPF-2S 1s -ward house 2P
 and unless you want to take me to your house
- (72) *ahla tō mēš-in mē ā, wa raft-ī ba mē*
 relatives 2s see:IRR-3P 1s SUB and go:PERF-2S with 1s
xāna
 marriage
 so that your relatives can look at me... since you have been married to me
- (73) *rōz-ē, di-rōz, bar mē xā šmā. ādī*
 day-a two-day carry:2S.IMPER 1s house 2P normal
č-um
 go:IMPF-1S
 a day or two, you must take me to your house. It’s tradition that I go
- (74) *wā tō, bāla ya'nī. tō tāt-ī xwā'-ī*
 with 2s up that_is_to_say 2s want:IMPF-2S sleep:IRR-2S
pana mē
 beside 1s
 up with you, that is to say. You want to sleep beside me
- (75) *naxa sōnty-ō, wa sā tu kardīd-ī mē inda*
 aboard raft-the and REPROACH drop:REAL-2S 1s inside
mušībit-ē!
 calamity-a
 on the raft, and now you! You have thrown me into calamity!
- (76) *mē nwāz č-um bāla qāniš wa gō ahla*
 1s tomorrow go:IMPF-1S up hunting and say:IRR relatives
tō
 2s
 [saying:] ‘Tomorrow I am going up hunting...’ and telling your relatives

- (77) *āmad-in ba mē: wōwōwō pis šēx-ō adafa būs-ē,*
 come:REAL-3P to 1s woe! son sheikh-the injuring
become:PERF-3S
 to come to me: [saying] ‘Woe! The sheikh’s son has been hurt!’
- (78) *mē fajja būr-um, sā ba rū mē jīr-in!*
 1s shocked become:REAL-1s now of face 1s see:REAL-3P
 I was shocked! Now they have seen my face!
- (79) *sīna mē jīr-in! mū-an mē jīr-in! ātiš-ō labaqa gid-ī*
 chest 1s see:REAL-3P hair-PL 1s see:REAL-3P fire-the igniting
do:REAL-2S
 They have seen my breasts! They have seen my hair! You have shamed me
 [literally: set my body on fire]!
- (80) *lāhm mē! damb gid-ī bā mē! čābē inča gid-ī*
 body 1s sin do:REAL-2S against 1s how? like_this
do:REAL-2S
 You have sinned against me! How could you have done this
- (81) *ba mē ā? mē škašt-um xō pi yā kār-ō.*
 to 1s INTERR 1s break:REAL-1S REFL from DEM
 work-the
 to me?! I want nothing more to do with this matter [literally: I have broken
 myself over this matter].”
- (82) *dgō ba yē jwān. sā bā yē na, samāha mē kin.*
 say:3S.IMPF to 3s good now all_right forgiving 1s
do:2S.IMPFR
 He said to her, “OK. Now, all right, forgive me.”
- (83) *br-im bāla xān mā. dgō ba yē br-im.*
 come:IMPER-1P up house 1P say:3S.IMPFR to 3s go:IMPER-1P
 Let’s go up to our house.” She said to him, “Let’s go.” (S787)

Abbreviations

ADVR	adverbializer	N	noun
AR	Arabic modern loanword gloss	NEG	negative
C	consonant	P	plural
COUNT	count marker (inanimate)	PERF	perfect verb form
DEM	demonstrative	PL	plural noun suffix
DV	deverb	REAL	realis verb form
EVID	evidential	REFL	reflexive pronoun
EX	existential	REP	reportive evidential
IMPER	imperative verb form	S	singular
IMPF	imperfect verb form	SENS	sensory evidential
INF	inferred evidential	SUB	subordinating enclitic
INTERR	question morpheme	TOP	topicaliser
IRR	irrealis verb form	V	short vowel
MIR	mirative verb form	VV	long vowel

References

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Anonby, Christina van der Wal. 2009. Words within words: The embedded poems of Kumzari fairy tales. Paper presented at the Conference *Documenting oral traditions in the non-Western world*, Leiden University, 27–29 August, 2009.
- Anonby, Christina van der Wal. 2014. Traces of Arabian in Kumzari. In Orhan Elmaz & Janet C.E. Watson (eds.), *Languages of Southern Arabia: Special session of the Seminar for Arabian Studies*, 137–146. Oxford: Archaeopress.
- Anonby, Christina van der Wal. 2015. *A Grammar of Kumzari: A mixed Perso-Arabian language of Oman*. Leiden: Leiden University dissertation.
- Anonby, Erik J. 2010. Kumzarītī [Kumzari alphabet chart]. Unpublished manuscript.
- Anonby, Erik J. 2011. Illustrations of the IPA: Kumzari. *Journal of the International Phonetic Association* 41(3). 375–380.
- Bashir, Elena. 2006. Evidentiality in South Asian languages. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the Lexical-Functional Grammar 2006 Conference*, 30–50. Palo Alto, U.S. A.: CSLI Publications.
- Bashir, Elena. 2010. Traces of mirativity in Shina. *Himalayan Linguistics* 9(2). 1–55.
- Bernabela, Roy S. 2011. *A phonology and morphology sketch of the Šihhi Arabic dialect of al-Ğēdih (Oman)*. Leiden: Leiden University, unpublished MA thesis.
- Caussin de Perceval, Armand P. 1853. Essai sur l'histoire des Arabes avant l'islamisme, pendant l'époque de Mahomet, et jusqu'à la reduction de toutes les tribus sous la loi musulmane. *The Calcutta Review* 21. 1–76. Calcutta: Sanders, Cones & co.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 1–41. Cambridge: Cambridge University Press.
- Dostal, Walter. 1972. The Shihuh of northern Oman: A contribution to cultural ecology. *The Geographical Journal* 138. 1–7.

- Eades, Domenyk. 2009. The Arabic dialect of a Šawāwī community of northern Oman. In Enam Al-Wer & Rudolf de Jong (eds.), *Arabic Dialectology: In honour of Clive Holes on the occasion of his sixtieth birthday*, 77–98. Leiden: Brill.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Complex Constructions (Language Typology and Syntactic Description, vol. 2)*, 1–51. New York: Cambridge University.
- Hawley, Donald. 1970. *The Trucial States*. London: Allen & Unwin.
- Holes, Clive D. 1990. *Gulf Arabic* (Croom Helm Descriptive Grammars). London: Routledge.
- Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In Talmi Givón (ed.), *Discourse and syntax (Syntax and Semantics, vol. 12)*, 213–241. London: Academic Press.
- Jayakar, A. S. G. 1902. The Shaheen dialect of Arabic. *Journal of the Bombay Branch of the Royal Asiatic Society* 4. 246–277.
- Johanson, Lars & Bo Utas (eds.). 2000. *Evidentials: Turkic, Iranian and neighbouring languages* (Empirical Approaches to Language Typology 24). Berlin & Boston: De Gruyter Mouton.
- King, Geoffrey R. 2001. The coming of Islam and the Islamic period in the UAE. In Ibrahim al Abed & Peter Hellyer (eds.), *United Arab Emirates: A new perspective*, 70–97. London: Trident Press.
- Kossmann, Maarten. 2000. *A Study of eastern Moroccan fairy tales* (Folklore Fellows' Communications 274). Helsinki: Suomalainen Tiedeakatemia Academia Scientiarum Fennica.
- Lancaster, William O. & Fidelity C. Lancaster. 2011. *Honour is in contentment: Life before oil in Ras al-Khaimah (UAE) and some neighbouring regions*. Berlin & Boston: De Gruyter.
- Lorimer, John G. 1915. *Gazetteer of the Persian Gulf, Oman, and Central Arabia*. Calcutta: Superintendent Government Printing.
- McWhorter, John H. 2007. *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford: Oxford University Press.
- Miles, Samuel B. 1994. *The countries and tribes of the Persian Gulf*. Reading, UK: Garnet Publishing.
- Nadjmabadi, Shahnaz. 1988. Identité ethnique contre nationalité: Le cas de l'île Larak (Golfe Persique). In Jean-Pierre Digard (ed.), *Le fait ethnique en Iran et en Afghanistan*, 65–74. Paris: Editions du CNRS.
- Neale, John Mason. 1873. *A History of the Holy Eastern Church*. London: Rivingtons.
- Palgrave, William Gifford. 1866. *Narrative of a Year's Journey through Central and Eastern Arabia* (1862–63), vol. 2, 3rd edn. London: MacMillan.
- Payne, Thomas Edward. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Potts, Daniel. 2012. *In the land of the Emirates: The archaeology and history of the UAE*. Abu Dhabi: Trident Press and Sultan bin Zayed's Culture and Media Centre.
- Propp, Vladimir. 1968. *Morfologija skazki* [The morphology of the folktale]. Austin, U.S.A.: American Folklore Society.
- Pustet, Regina. 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- al-Rawas, Isam. 2000. *Oman in early Islamic history*. Reading, UK: Garnet & Ithaca Press.

- Ross, Edward C. 1874. Annals of Oman. *Journal of the Asiatic Society of Bengal* 43. 111–196.
- Rubin, Aaron D. 2010. *The Mehri language of Oman*. Leiden: Brill.
- Simeone-Senelle, Marie-Claude. 1997. Modern South Arabian languages. In Robert Hetzron (ed.), *The Semitic languages*, 378–423. London: Routledge.
- Skjærvø, Prods Oktor. 1989. Languages of southeast Iran: Lārestānī, Kumzārī, Baškardī. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 363–369. Wiesbaden: Reichart.
- Stilo, Donald. 2004. Coordination in three Western Iranian languages: Vafsi, Persian and Gilaki. In Martin Haspelmath (ed.), *Coordinating constructions*, 269–330. Amsterdam & Philadelphia: Benjamins.
- Thomas, Bertram. 1930. The Kumzari dialect of the Shihuh tribe (Musandam), Arabia, and a vocabulary. *Journal of the Royal Asiatic Society* 62(4). 785–854.
- Ulrich, Brian John. 2008. *Constructing Al-Azd: Tribal identity and society in the Early Islamic centuries*. Ann Arbor, U.S.A.: ProQuest.
- Utas, Bo. 2006. ‘Genres’ in Persian literature 900–1900. In Gunilla Lindberg-Wada (ed.), *Literary history: Towards a global perspective*, vol. 2, 199–241. Berlin & Boston: De Gruyter.
- Waltisberg, Michael. 2011. Syntactic typology of Semitic. In Stefan Weninger (ed.), *The Semitic languages: An international handbook*, 303–329. Berlin & Boston: De Gruyter Mouton.
- Watson, Janet C.E. & Domenyk Eades. 2012. *Documentation and ethnolinguistic analysis of the Modern South Arabian languages*. Leverhulme Trust Project Grant outline. http://www.salford.ac.uk/_data/assets/pdf_file/0012/173010/MSALProjectInformation.pdf (accessed 16 February 2013).
- Watson, Janet C.E. 2012. Dialects of the Arabian Peninsula. In Stefan Weninger (ed.), *The Semitic Languages: An international handbook*, 897–908. Berlin & Boston: De Gruyter Mouton.
- Zwemer, Samuel Marinus. 1902. Three journeys in northern Oman. *The Geographical Journal* 19(1). 54–64.

5. The Caspian region and south Azerbaijan: Caspian and Tatic

Donald L. Stilo

1. Introduction

This chapter focuses on two Northwest Iranian (NWI) families to the south and west of the Caspian Sea, Caspian and Tatic. The internal relationships of these languages remains a topic of controversy; the presentation adopted here follows Stilo (1981). Much of the data stems from fieldwork undertaken in the region in the 1970's, but supplemented by more recent published data. The Caspian group represents one long, unbroken dialect chain extending some 500 km from the Mazanderani (Māzanderāni) dialects of the Gorgān area in the east to the Gilaki dialects of Anzali in the west. Within Caspian, a major division occurs between Eastern Gilaki (EG) and Western Gilaki (WG), whose boundary is formed by the Sefidrud river. WG includes the dialects of Rasht, Fuman, Anzali, and EG includes Lāhijāni, Langerudi, Māchiāni, among others. In addition, I consider the transitional dialects roughly of the area from Chālus to Rāmsar to be a third "Central Caspian" language.

The Tatic family consists of: A) highly diverse Tati (Tāti) dialects that are sparsely spread over a discontinuous area extending from Vafsi in the south near Saveh to Kilit (now extinct) of Nakhichevan/Naxjavān province of the Republic of Azerbaijan north of the Araxes; B) Talyshi (Tāleshi/Taleshi in Iranian sources), ranging from the northernmost dialect of the Masally area of Azerbaijan to the Rudbār valley in Iran; and C) "Tatoid", two Tati-like offshoots, Rudbāri and Tāleqāni/Alamuti, as discussed below. Also note the usage of "*Talysh people*"/"*the Talysh*" vs. "*the Talyshi language*".

These sub-groups and varieties are summarized in the following list. The numbers in parentheses refer to the numbering in Fig. 1 below:

Caspian
Mazanderani (1)
Central Caspian (2)
 Ramsar, Tonekabon, Kelardasht
Eastern Gilaki (3)
 Lahijan, Langerud, Machian
Western Gilaki (4)
 Rasht, Fuman, Anzali

Tatic

Tati

S. Tati (7)

C. Tati (8)

N. Tati (9)

Talyshi

S. Talyshi (10)

C. Talyshi (11)

C.-N. transitional (12)

N. Talyshi (13)

Tatoid

Taleqāni-Alamut Tatoid (5)

Rudbār Tatoid (6)

In this chapter, I have singled out four varieties to represent the four main groups discussed here for more detailed treatment: Sāravi (Mazanderani), Lāhijāni (Gila-ki), Leriki (N. Talyshi), and Vafsi (Tati), but other, especially undocumented, varieties will be introduced when space permits.

1.1. Geographical Situation

The Caspian littoral, due to the presence of several of Iran's major rivers and their tributaries as well as to its relatively low elevation blocked off from the plateau by the high physical barriers provided by the Alborz and Talesh mountain ranges, has as a result a quite humid, mild, subtropical climate and is known for its dense forests, rice paddies, tea plantations, citrus groves, fisheries, and diverse fauna. Due to the climate and fertile land, Caspian and Talyshi groups are spoken in densely populated areas.

The Caspian languages are located along the southern Caspian littoral extending eastward almost to Gorgān. Talyshi is located along the southwest and lower areas of the western Caspian littoral, especially in the cities of Māsule, Māsāl, Hashtpar/ Asālem in Iran, extending northwards into the Republic of Azerbaijan through the major cities of Astara, Lenkoran, and Masally. Varieties of Caspian (e. g., Kelārdashti, Elāshti, Velātrui) and Talyshi (esp. Māsulei, Anbarāni, Leriki) are spoken higher up in the adjacent verdant mountain slopes and valleys facing the Caspian. Mazanderani has spread as far south as Ure between Demāvand and Āb-e Ali.

There is, however, a complete linguistic break with no mutual intelligibility between Caspian and Talyshi, specifically neighboring Western Gilaki and Southern Talyshi. They belong to two different subgroups of NWI that arrived in the area from different directions and have no transitional forms as a bridge between them.

By contrast to Talyshi, Tati languages are located on the Iranian plateau from

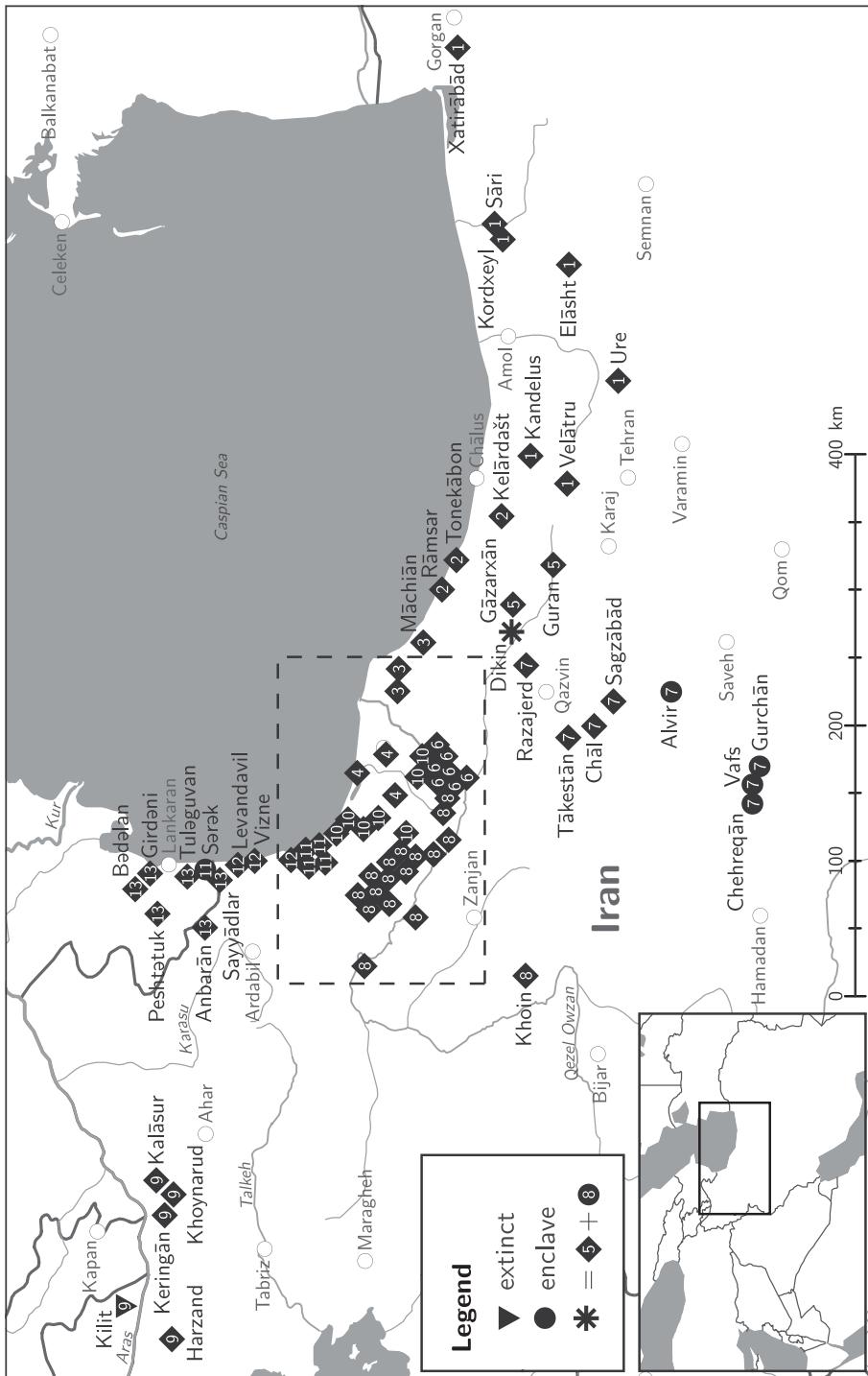
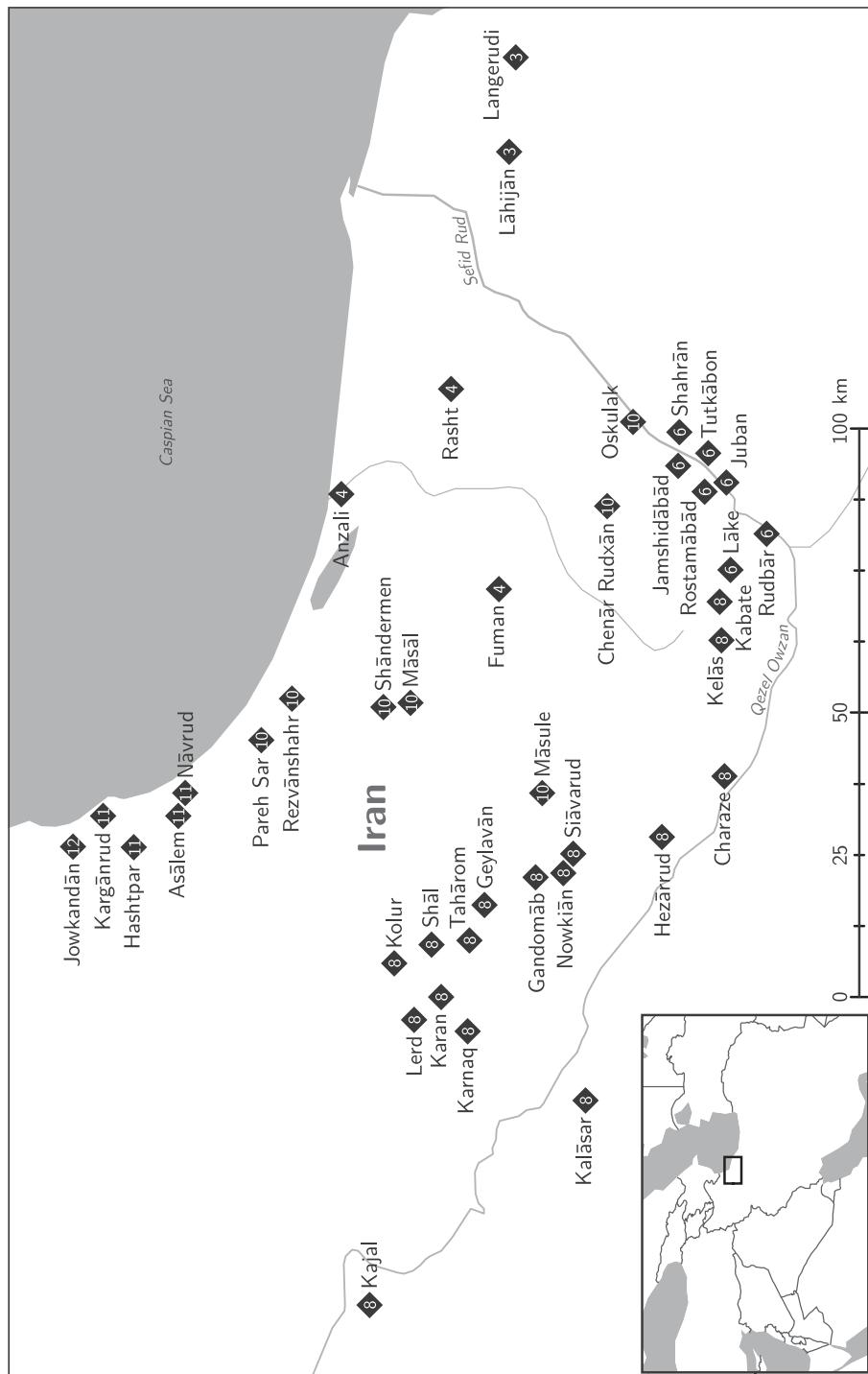


Figure 1: Locations of Caspian and Tatic varieties



the Araxes in the north and southwards to Saveh and Arak, and are sparsely distributed among villages that generally only support small populations. Notable exceptions are the Tākestān area, Kolur, and Vafs. At least half of the populations speaking the latter two are located in Tehran, although they retain regular contact with their villages.

It should also be noted that in the last 40–50 years, the major cities along the Caspian coast in the Talyshi language zone, especially in Iran, have become predominately Azerbaijani Turkish speaking (due to in-migration). Today we find the majority of Talyshi speakers in villages outside the major cities.

2. The internal relationships of Caspian and Tatic

2.1. Caspian

Claims are often made for Galeshi as a separate Caspian dialect. While they may be considered by some (and themselves) to be a distinct ethnic group, certainly based on a distinct life-style, the linguistic data do not support any putative claims for a single unified dialect. The Galesh are cowherds, whence the word *Gāleš* < **gawa-raxšaka-* ‘the protector of cows’ (Asatrian and Borjian 2005: 50), in mountainous areas extending along the Alborz from at least the Sefidrud through to eastern Mazanderan. Varieties of Galeshi differ from point to point just as the dialects of the lowlands do. In fact, any given Galeshi dialect is probably much closer to the corresponding lowland dialect to the north than it is to another Galeshi dialect at a similar east-west distance.

This connection makes perfect sense given that until recently the pastoral population probably led a summer camp/winter camp (*yeylaq/qishlaq*) transhumant life-style, migrating down to the closest lowland point in the winter season. This caused their “outside” focus to turn to the lowland area, rather than to other Galesh groups to the east or west. While Galesh dialects across most of Gilan may be mutually intelligible, they clearly exhibit some differences from one another. The Galeshi of Gilan is totally different from the Galeshi of Mazanderan (see Borjian and Borjian 2007a, 2007b, 2008). Further discussion of Galeshi will have to remain for a separate investigation.

2.2. The Tatic family

Tatic consists of two closely related main groups, Tati and Talyshi, both of which can be further grouped into Northern, Central, and Southern clusters. While the Talysh population is quite dense and forms a classic dialect chain extending southwards from (Azerbaijan) Masally, Lenkoran, Lerik, Āstārā, (Iran) Anbarān, Qal’eh, Sayyādlar, Levandavil; Vizne, Jowkandān; Kargānrud, Asālem, Nāvrud;

Paresar, Tāleš-Dolab; Xošābar; Shāndermen, Māsāl, Māsule, culminating in the easternmost (as far as I know) Oskulak in the Rudbār valley and southernmost Chenār Rudxān (Bazin 1981: 113).

Caucasian Tat, a Southwest Iranian (SWI) language resettled from southern Iran in late Sassanian times, is not associated in any way with the Tati family or the term Tatic. Its closest relative is Persian and the similarity in names is purely coincidental.

The Tati languages by contrast are disjointed and dispersed, as follows:

- Southern Tati with nine dialects in the Tākestān area of Qazvin (Yarshater 1969), Chāli, Tākestāni, Sagzābādi, etc. and to the east, Eshtehārdi. Razajerdi merits special status as a transition from Southern Tati to Tāleqāni (Tatoid).
- After a break from Southern Tati, Central Tati picks up to the northwest: (1) the Tāromi group (Yarshater 1970)—Hezārrudi, Kalāsari, Bākoluri, Charazei, etc. of Upper Tārom in Zanjān province—and including Kelāsi and Kafteji in Gilan province; (2) Koluri, Shāli, etc. in Shāhrud district (*bæxš*) of Khalkhāl, Ardabil province (Yarshater 1959); (3) Gandomābi, Siāvarudi, Nowkiāni, etc.; (4) Kahali, Geylavāni, Tahāromi, etc.; (5) Karani, Lerdi, Dizi, Karnaqi; and (6) slightly to the west, Kajali (Yarshater 1960) of Kāqazkonān district, Miāneh county, showing some transitional lexical features to Northern Tati.
- The first five groups above form a spectrum-like dialect chain but seasonal work and trade bring them into regular contact with the Southern Talysh of Māsule, Māsāl, and Shāndermen, creating a conduit for lexical and grammatical borrowing, probably in both directions. The Kajalis' Sunni denomination also brings them into similar contact with the Sunnis of the Central Talyshi areas.
- To the north of Central Tati, after a substantial break, we find disparate clusters of Northern Tati dialects along the Aras (Araxes) river: Kalāsuri-Khoynarudi, Keringāni, Harzani, and (extinct) Kilit (on the north side of the Araxes, near Ordubad, Nakhjavan province, Republic of Azerbaijan).

2.2.1. Notes on specific enclave speech varieties and their communities

Certain languages are the results of population movements that brought speakers of a few Tati or Talyshi varieties to other areas within Tatic, to the periphery of this zone, or outside it. In this chapter, I will use the term “enclaves” for these isolates.

Talyshi-speaking groups

Charozh is a type of Central Talyshi (<*čarāj* ‘a Talyshi tribe of Iran’ (‘Abdoli Koluri 1992: 43)), spoken by some 100–200 people in Sarak and Degadi in the Astara Talyshi-speaking area in the Republic of Azerbaijan. As Sunnis who felt some persecution in Shiite Iran, they were invited by Tsarist Russia to resettle north of the border soon after the 1828 Treaty of Turkmanchay. While Charozh has

retained many emblematic features of C. Talyshi, it also displays heavy influence from the surrounding Northern Talyshi.

Oskulak, a village in the Rudbār valley with some 1000 inhabitants at the time of my 1976 visit, is evenly divided between two linguistic groups: Oskulaki-Talyshi and Oskulaki-Rudbāri (Tatoid, §2.3). The former group is quadrilingual (Talyshi, Rudbāri, Gilaki, Persian) and the latter trilingual (Rudbāri, Gilaki, Persian).

Tati-speaking groups

Dikin, a village of 850 inhabitants at the time of my visit in 1976, located to the northeast of Qazvin city in the Alamut valley (Stilo 2018b), has three different groups reflecting a divide between: A) adherents of a secret sect found in the Alamut area who call themselves “Marāqeи” or “Kalle-bozi (Goathead)”, who speak Dikin-Marāqeи (ca. 500 people), and the regular Shiite villagers (called “Pashei (Mosquito)” by the Marāqeis), who speak either B) Dikin-Pashei, a typical Alamuti (Tatoid) dialect of the area, or C) standard Tehrani Persian as a first language. Marāqeи groups are found in some 16 villages in the area where they only coexist with non-Marāqeи Shiites and have no villages exclusively of their own. Marāqeи dialects are generally of a highly conservative Tati type (§4.2), very different and unintelligible to the Tatoid “Pashei”.

Vafsi (Stilo 2004a), spoken in four villages (the Vafs-Chehreqān-Fark cluster and Gurchān, somewhat removed), has a seemingly very complex history. It is often classified with either Tati (Stilo 1981), the Central Dialect group (Lecoq 1989), or as transitional between them (Redard 1970). In my opinion, it reflects a Tatic admixture from farther north overlaying a Central Dialect substratum, as seen in certain shared innovations typical of N. Tati, not found in Southern Tati. There is an additional stratum of a Kurdish-speaking group, yet to be determined. This area has attracted many different groups, as it is a highly fertile, high mountainous plain endowed with a wide variety of orchards that extend for kilometers around the villages and is especially attractive for its cooler weather in the summer months. My views on this complex layering in the history of Vafsi is a topic for future publication.

Kafteji and Kelāsi, spoken in Kafti and Kelās (Pers. Kabate and Kalās), are located just within Rudbār district of Gilan province and are the easternmost of the Tāromi cluster. In the early 1970s the villages, located some 12 km apart with no intervening villages, consisted of about 350 inhabitants each. With the socio-economic upheavals of the 1970s, however, almost all the villagers left for larger cities and industrial centers.

Table 2A. Some common lexemes across the Caspian continuum

FEATURE	Rashtí	Lāhijānī	Galeshi	Langerudi	Rāmsari	Tonkāboni	Kelārdāšt ¹	Kandelusí ²	Sāraví	Velātrui	Urei	Elāšti	Xatirabdi
crow	<i>kəlač</i> , <i>kəlaq</i>		<i>kəlač</i> , <i>kəlač</i>	<i>kəlač</i>		<i>kəlač</i> , <i>qəla</i>		<i>kəlač</i>	<i>kəlač</i> , <i>kəlaq</i>			<i>kəlač</i>	<i>kəlač</i>
throw, drop-PT	<i>i-/də-gad,</i> <i>ta-vəda</i>	<i>to-da</i>	<i>gent,</i> <i>gəne,</i> <i>tua-da</i>	<i>də-(r)</i>	<i>tae-væ-d</i>	<i>ingæt</i>	<i>enget</i>	<i>aengen</i> (pres.)	<i>d-i(n)gu</i>	<i>engess</i>	<i>aengən*</i>	<i>d-engo</i>	
boy	<i>rey</i> , <i>re(e)k</i>	<i>pəsər</i>	<i>rikə</i>	<i>væče,</i> <i>reka</i>	<i>væče,</i> <i>reka</i>	<i>rika</i>	<i>rika</i>	<i>ri:ko</i>	<i>rika</i>	<i>rika</i>	<i>veče</i>		
sister	<i>xaxur</i>	<i>xaxor</i>	<i>xaxor</i>	<i>xaxor</i>	<i>xaxor</i>	<i>xaxor</i>	<i>xaxor</i>	<i>xaxer</i>	<i>xaxer</i>	<i>xaxer</i>	<i>xaxer</i>	<i>xaxer</i>	<i>xaxer</i>
buy-PR	<i>hin</i>	<i>hen</i>	<i>hin</i>	<i>hi(n)</i>	<i>hin</i>	<i>xər</i>		<i>xərin</i>		<i>xər</i>	<i>xrin</i>	<i>xər</i>	
burn-PR	<i>suj</i>	<i>suj</i>	<i>suj</i>	<i>suj</i>	<i>suj</i>	<i>suj</i>		<i>suz</i>			<i>suj</i>		
cook-PR	<i>pəz, pəč</i>	<i>pəč</i>	<i>puč</i>	<i>puč, poč</i>		<i>pəč</i>	<i>pəjən</i>	<i>pəj, pej</i>		<i>pəj,</i>	<i>pəj</i>		
louse	<i>səbəj</i>	<i>subuj</i>	<i>subuj</i>	<i>səbəj,</i> <i>osboj</i>		<i>espej</i>	<i>esbij</i>	<i>esbij</i>					
nose	<i>vini</i>	<i>domay</i>	<i>vini</i>	<i>vini</i>	<i>fini</i>	<i>veni</i>	<i>feni</i>	<i>vəni</i>	<i>veni</i>	<i>feni</i>	<i>demay</i>		

*Forms unavailable for Elāšti (Pahlavān 2003) were supplied by Yuši data (Tāhhāz 1963), another mountainous dialect

¹ Kelārdāšt material is from Kalbāsi 1376/1997

² Kandelusí material is from Jahāngiri, 'AliAsghar 1367/1988

2.3. Comparative lexicon of Caspian and Tatic

2.3.1. Caspian lexical features

Tables 2A–C present various aspects of the lexical composition and distribution across the Caspian continuum. Examples are restricted here to only a few in each category. As Table 2A shows, many lexical items distributed across the Caspian family still have their expected NWI reflexes and/or some special Caspian feature.

The Caspian family has had intimate contact with Persian over a rather long period as is evident in both lexical and grammatical domains. It is my opinion that a certain number of Persian lexical borrowings are from a much earlier stage of Caspian prehistory since they are very consistently distributed throughout, and are completely integrated into, the Caspian family as a whole, e. g., Lāhijāni: *don/donəst* ‘know’, *zən/ ze* ‘hit’, *gu/gut* ‘say’, *zənəy* ‘woman’, *səg* ‘dog’, *mayi* ‘fish’, *atəš* ‘fire’, *gul* ‘flower’, *asiyow* ‘mill’, *fərde* ‘tomorrow’, *diruz* ‘yesterday’, *bəhar* ‘spring’, *su* ‘three’.

By contrast, Table 2B shows that numerous later Persian borrowings are neither universal throughout the family nor integrated into the more conservative dialects. The original NWI-type word is still extant in peripheral and even urban areas. Table 2C presents a small sampling of the distribution of lexical items serving as areal “links” along the Caspian dialect chain. Such a study must perforce allow for gaps in our knowledge of these languages for undocumented synonyms that would be included if the data were present.

2.3.2. Tatic lexical features

Although Tati and Talyshi are very closely related, there is wide variety within each group—much wider than that of the Caspian group. As I have shown elsewhere (Stilo 2015), Northern and Southern Talyshi should be considered two independent languages. The diversity within Tati is even more striking, with Northern Tati being an extremely unusual case within West Iranian overall.

As a counterpoint to this statement about diversity within Tati and Talyshi, Table 2D briefly shows that Northern Talyshi is clearly more closely related to Northern Tati (rows 1–6) than it is to Southern Talyshi. Central Talyshi has close ties to the more northerly Central Tati dialects (rows 7–9), and Southern Talyshi is quite closely related to other forms of Central Tati (rows 10–12).

2.4. Tatoid groups: a brief statement

The Rudbāri group, with many mutually intelligible dialects, is located in the Rudbār valley on both sides of the Sefidrud river extending some 35 km from

Table 2B: Some Persian loanwords from a later period

	Rashti	Lâhjan	Galeši	Langrudi	Mâč ³ Râmsar	Tonkab	Kelrdšt	Sâravi	Urei	Elašti	Kordxl Xatirb
down (jir)	jir	jir	jir	jir	jir	jir	jir	zir		jer	zer
hungry	vîšta	gušnə	vâšnə	gursənəy	vêšna	vêšna	vêšna	vêšna	vêšna	vêšna	gašn̩l
kidney		fak		væk	volv	qolhe	qolbe			qolve	
pull out, take off (clothes)	birun avər	kən	kən		kən	kən	boj			kən	
seed	tuxm	tuxm	tuxm, tum	tuxm	toxm, tum	toxm, tum	tim	tim	tim	tim, toxm	
sew-pr	duz	duz	duy	duy	duij, duz		duij	duij	duij	duij	duz
snow	berf, vərf	berf, vərf	vərf	bərf, vərf	vərf, bərf	bərf	vərf, bərf	bərf	vərf	vərf	vərf
son-in-law	damad	zome	damad	zoma zoma	zoma	zoma	domad	domad	domad	domad	damad
tear (drop)	ašk	ašk		osiru	osure	æseri	æsli			æsli	
up (jor)	jor	jor	jəwr	jor	jor	(bu)jor	jor, jor	jor	bala	yor	bala
weave, knit	baf	baf		vaj	aj	vaf	b-vof				
wolf	gurg	gurg		gurg	vərg	vərg	vərg	vərg	vərg	vərg	vərg

³ Mâčiāni material is from Farzpur-Mâčiāni 1964, 1965

Table 2C: Intra-Caspian lexical isoglosses

Item	Rashki	Lăhijāni	Galeshi	Langrud	Măčān	Rāmsari	Tonkbn	Kelārdš	Kandlus	Săravi	Velātrui	Ure	Elašti	Xatirbd
1 (1, 5)	sparrow <i>čičini</i>	čušnək	<i>məljə</i>	<i>čiš(ə)nək</i>	<i>məljəɛ'</i>	<i>milje</i>	<i>meļje</i>	<i>mička</i>	<i>mička~</i>	<i>miška</i>	<i>mička</i>	<i>mička³</i>	<i>jika</i>	
2	child	<i>zaj, zaak zak,</i> kočtay	<i>zaak</i>	<i>zæk/zak</i>	<i>væčæ</i>	<i>væčæ</i>	<i>væč(č)e</i>	<i>væčæ</i>	<i>væčæ</i>	<i>væčæ</i>	<i>væčæ</i>	<i>væčæ</i>	<i>væčæ</i>	
3	large, big (4, 12)	<i>pile,</i> <i>pil(l)ə</i>	<i>pila,</i> <i>pildanl</i>	<i>pilla</i>	<i>pile, pilla</i>	<i>pilači</i>	<i>gæt</i>	<i>gæt, kæl gæta,</i> <i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>get</i>
4	throw, drop-pf	<i>ta-vəda,</i> <i>i-də-gad</i>	<i>to-da</i>	<i>gent,</i> <i>tua-da,</i> <i>də- (r)</i>	<i>əgen¹</i>	<i>tæ-væ-d</i>	<i>inget</i>	<i>enget</i>	<i>əngen</i>	<i>d-i(n)gu</i>	<i>engess</i>	<i>d-ingu</i>	<i>d-engo</i>	<i>'put'</i>
5 (1, 5)	sparrow <i>čičini</i>	čušnək	<i>məljə</i>	<i>čiš(ə)nək</i>	<i>məljəɛ'</i>	<i>milje</i>	<i>meļje</i>	<i>mička</i>	<i>mička~</i>	<i>miška</i>	<i>mička</i>	<i>mička¹</i>	<i>jika</i>	
6 14)	ant (6, <i>putur</i>	<i>mur(čə),</i> <i>pitar</i>	<i>pitar</i>	<i>pitar</i>	<i>potor</i>	<i>putar</i>	<i>melijə</i>	<i>mejlə</i>	<i>melijə</i>	<i>mejile</i>	<i>melije</i>	<i>melije</i>	<i>melijl</i>	
7	cat (7, 17)	<i>piča</i>	<i>pičə</i>	<i>pičə</i>	<i>piča</i>	<i>pučað</i>	<i>piča</i>	<i>piča</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	
8	hole	<i>sulax,</i> <i>xulə</i>	<i>sulax</i>	<i>luk,</i> <i>sulay</i>	<i>luk,</i> <i>sulax</i>	<i>luka,</i> <i>surəx</i>	<i>luka,</i> <i>surəx</i>	<i>veher,</i> <i>loxar,</i> <i>lex</i>	<i>le:x,</i> <i>ve:he:r</i>	<i>sulax</i>	<i>li</i>	<i>selax</i>		
9	spider	<i>labdon</i>	<i>labdon</i>	<i>labdon</i>	<i>kartələm</i>	<i>kartən</i>	<i>vennali</i>	<i>baend</i>	<i>kartənək²</i>					
10	lip	<i>lab</i>	<i>lab</i>	<i>məč</i>	<i>mači</i>	<i>tok</i>	<i>tok</i>	<i>tek, lošte tek, lo:</i>	<i>tek, lo,</i>	<i>lu:šəe</i>	<i>tek,</i>	<i>luš(ə)e</i>		

Item	Rashči	Lāhičāni	Galešhi	Langrud	Mäčiān	Rāmsari	Tonkbn	Ketārdš	Kandlus	Sāravī	Velātrui	Ure	Elāšti	Xatirbd
11 girl	<i>kor(ta)</i>	<i>laku</i>	<i>laku</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>kija</i>	<i>dēter</i>	
12 large, big (4, 12)	<i>pile,</i> <i>pil(θ)ə</i>	<i>pila,</i> <i>pildanl</i>	<i>pile/ð,</i> <i>pilowl/</i>	<i>pilla,</i> <i>pilæči</i>	<i>gætə</i>	<i>gæt</i>	<i>get, kæl gæta,</i> <i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>gæt</i>	<i>gæt,</i> <i>gæti</i>	<i>gæt,</i> <i>gæti</i>		
13 broom	<i>ja:ru</i>	<i>jaru</i>	<i>jaru</i>	<i>taru</i>	<i>toru</i>	<i>saje</i>	<i>saje</i>	<i>sajə</i>	<i>saze</i>	<i>sɔ:jɛ</i>	<i>sazə</i>	<i>sazə</i>	<i>sazə²</i>	
14 ant (6, 14)	<i>mur(čɔ),</i> <i>putur</i>	<i>pitar</i>	<i>pitar</i>	<i>potor</i>	<i>putar</i>	<i>mælje</i>	<i>mejlæ</i>	<i>melje</i>	<i>mejile</i>	<i>mejile</i>	<i>mejile</i>	<i>mejile</i>	<i>mejili</i>	
15 come!	<i>bia</i>	<i>be</i>	<i>bia</i>	<i>bia</i>		<i>boro</i>	<i>(y)a</i>	<i>beru</i>	<i>beru~b̥oro</i>	<i>beru</i>	<i>beru</i>	<i>beru</i>	<i>bia</i>	
16 self	<i>xu~xo,</i> <i>xud-</i>	<i>xu</i>	<i>xo, xu,</i> <i>xud</i>	<i>xošte(rə)</i>	<i>xoštere,</i> <i>xod-</i>	<i>šer</i>	<i>še:</i>	<i>še, xod</i>	<i>xæ, xe,</i> <i>xod-</i>	<i>še,</i> <i>xed-</i>	<i>še¹</i>	<i>še¹</i>	<i>xaš,</i> <i>xad-</i>	
17 cat (7, 17)	<i>piča</i>	<i>pičə</i>	<i>pičə</i>	<i>piüča</i>	<i>puča/ð</i>	<i>piča</i>	<i>piča</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	<i>bameši</i>	

1) Forms unavailable for Mäčiāni were supplied by Väjärgähi data (Amirzädeh-Väjärgäh 1375/1996), some 10 km to the east of Mäčiān.

2) Forms unavailable for Elāšti were supplied by Yuši data, another mountainous dialect.

3) Forms unavailable for Velātrui were filled in by Shahpoli in the same region.

Table 2D

Gloss	Northern Tati	Northern Talyshî	C. Talyshi	Southern Talyshî	Central Tati
Harzani ⁴	Keringān ⁵	Kalâsuri N.Talyshî	Viznei	Jowkand	Mâsâli
1 lip	lev	löbüüd	lipüt, lïv	lev	lëb, laſe
2 stomach	ged	qæðæ	gæðæ lævæ	laevæ	ter næ laevæ
3 neck	giri	geri	giy	mał'	gærðæn gærðæn
4 moon	öšmæ,	ušma,	ušmæ	ošum	mæng ma
5 break- pt/intr	sısd	sıst	si, pe-si	e-(š)ši iši	(a)-čækəst čækæ
6 laugh- pres	sər, sir	ser	sir	xur	xendæ xænnæ k k
7 neck	giri	geri	giy	mał'	gærðæn gærðæn
8 small, little	vede(k)	giidæ	gæðæ riłk	goč, geč	ruk villæ geč
9 laugh- pres	sər, sir	ser	sir	ser	xendæ xænnæ k k
10 frog	qarbaqa	bæzæuzq	væzæx	vezzæk	verzeek merəñjuncæ
11 elbow		niven- jaenæ		verzeek əšküb	gužga ušküf eškof
12 stomach	ged	qæðæ	gæðæ lævæ	laevæ ter næ	ter næ laevæ

1) A velarized /l/, sometimes also called ‘dark’ or ‘back’ /l/.

⁴ Harzani material is from Zoka 1336/1957⁵ Keringān material is from Zoka 1332/1954

Pan-Tatic lexicon

The following is a small sampling of lexical items that all (or almost all) Tatic languages, not including Tatoid, share (also found outside Tatic):

Table 2E

Item	N.Talyši Asālemi	Mäsulei Harzani Käläsur Kajali	Koluri	Karani	Gandomäb Hezärrud Kafteji Dik.-Maräq Chäli	Alviri	Väfsi
daughter <i>kine</i>	<i>kełtae</i>	<i>kellæ</i>	<i>kineæ</i>	<i>détae</i>	<i>det</i>	<i>détae</i>	<i>kina</i>
girl <i>kinæ</i>	<i>kełtæ</i>	<i>kellaæ</i>	<i>kinæ</i>	<i>killig</i>	<i>kelle</i>	<i>killegæ</i>	<i>kelkæ</i>
dog <i>sipæ</i>	<i>esbæ</i>	<i>əsba</i>	<i>esbe</i>	<i>espæ</i>	<i>sebæ</i>	<i>sbæ</i>	<i>æsbaæ</i>
flower <i>vil</i>	<i>vat</i>	<i>gal</i>	<i>vel</i>	<i>vel</i>	<i>vel</i>	<i>vel*</i>	<i>vélæ</i>
hole <i>hil</i>	<i>xəl</i>	<i>xəl</i>	<i>henæ</i>	<i>hełce</i>	<i>xel</i>	<i>velæ</i>	<i>velæ</i>
kidney <i>velk</i>	<i>velk</i>	<i>verek</i>	<i>qolba</i>	<i>vek</i>	<i>væk</i>	<i>vék(k)æ</i>	<i>vékæ</i>
shovel <i>hi(j)æ</i>	<i>xeyæ</i>	<i>xieæ</i>	<i>hiyæ</i>	<i>xeyæ</i>	<i>xieæ</i>	<i>xiyæ</i>	<i>xiyæ</i>
burn-pr <i>væš</i>	<i>væš</i>	<i>væš</i>	<i>væš</i>	<i>væš</i>	<i>væš</i>	<i>væši</i>	<i>de-væsen</i>
jump-pr <i>væšt</i>	<i>væz</i>	<i>nev</i>	<i>væz</i>	<i>daæ-væz</i>	<i>væz</i>	<i>vaz k.</i>	<i>væz</i>
jump-pt <i>væšt</i>	<i>væšt</i>	<i>nevæst</i>	<i>væšt</i>	<i>daæ-væšt</i>	<i>væšt</i>	<i>væſt</i>	<i>ö-væſt</i>
stand-pr <i>indin</i>	<i>mon</i>	<i>vanden*</i>	<i>vendæ</i>	<i>vendar</i>	<i>vendar</i>	<i>iſt</i>	<i>a-vendar</i>
						<i>is</i>	<i>har-vender</i>
*† = velar / *Mässali				*'run'	*'bud'		

Rudbār city in the south to the northernmost Oskulak. The fertile Rudbār valley is known for its olive groves and rice paddies. The dialects that I was able to identify and record during multiple trips to the area in 1976 consist of: (west bank, Sefidrud) Rudbāri; Lakei; the Juban cluster (Jubani, Sarāmarzi, Bāboni, and others); Rostamābādi (formerly Kalurazi); Jamshidābādi; Oskulaki; (east bank) Shahrāni, Tutkāboni. To this list we can add the dialect of Dogowharān as presented in Lazard (1990).

The other Tatoid group, Tāleqāni-Alamuti, consists of a dialect continuum spoken in an extensive area more densely populated than most plateau village areas. Its North-South extension extends from the highest Alborz villages to the Karaj-Abyek areas south of the local highway. Its East-West breadth ranges between the two major routes leading from the plateau through the Chālus and Rudbār passes to the Caspian. The western dialects, along the Shāhrud river, to my knowledge, have never been documented and it is not clear exactly how far they extend and where Rudbāri and Gilaki take over. The area is also interspersed with numerous Azeri and Kurmanji-speaking villages.

The only groups identified with labels until now are Tāleqāni and Alamuti, spoken along the parallel and adjacent Tāleqān and Alamut rivers northeast of Qazvin. The confluence of these two rivers forms the larger Shahrud river which in turn helps form the Sefidrud. For convenience sake, I shall refer to this linguistic continuum henceforth in this chapter as Tāleqāni with the understanding that it also includes Alamuti.

Tāleqāni and Rudbāri merit a special status that I have called “Tatoid”. Both groups, in my opinion, were originally Tati languages that lost all the structural hallmarks of Tati under heavy Caspian and Persian influence. The only traces, albeit not many, that hint at Tati origins are in the lexical domain, as seen in the brief sample in Table 2F. Table 2F also hints at a larger study of lexical items connecting Tāleqāni and Rudbāri to the Tatic and Caspian languages that are geographically closest to each of them — Kelārdashti and Chāli for Tāleqāni-Alamuti; Kafteji/Kelāsi and Rashti for Rudbāri.

Table 2F: Lexical similarities between Tatoid groups and the Caspian families

GILAKI		C. TATI	RUDBĀRI (TATOID)	S. TATI	TĀLEQĀNI (TATOID)	C. CASP.
LEXEME	RASHI	KAFTEJIR-(KELĀSI)	TUTKĀBONI	JUBANI	CHĀLI	ALAMUT AREA ⁶
ant	<i>putur</i>	<i>murčanæ</i>	<i>putir</i>	<i>moronj</i>	<i>mörčöne</i>	<i>murčane</i>
cook-PR	<i>pæč</i>	<i>pæč</i>	<i>pæč</i>	<i>pej, peč</i>	<i>pæč</i>	<i>pæč</i>
crow	<i>kəlač</i>	<i>qælač</i>	<i>kelač</i>	<i>kelač</i>	<i>qolač</i>	<i>kelač</i>
cry-PR	<i>gəryə (N)</i>	<i>bərəm*</i>	<i>gerye (N)</i>	<i>(be)rmæm</i>	<i>goræ (N)</i>	<i>borme (N)</i>
down	<i>jir</i>	<i>jir; jer</i>	<i>jer</i>	<i>jær; jaṛa</i>	<i>jir</i>	<i>jir</i>
fall-PR	<i>dæ-kæf</i>	<i>gen</i>	<i>dæ-k(ah)</i>	<i>gæn</i>	<i>k-</i>	<i>kow</i>
fall-PT	<i>dæ-kæft</i>	<i>kæt</i>	<i>dæ-kæt</i>	<i>kæt</i>	<i>ke(t)</i>	<i>kæt</i>
fox	<i>ruba</i>	<i>los</i>	<i>rubah</i>	<i>luas</i>	<i>luas</i>	<i>rua</i>
girl	<i>kor</i>	<i>keléktæ</i>	<i>kilkay(y)</i>	<i>kelka</i>	<i>teti</i>	<i>dutær</i>
kidney	<i>fæk*</i>	<i>vækæ</i>		<i>væk*</i>	<i>væk, vurdū</i>	<i>qolve</i>
large	<i>pile</i>	<i>pillæ</i>	<i>pila</i>	<i>pile</i>	<i>pil(l)æ</i>	<i>gæt*</i>
louse	<i>sabj</i>	<i>sbj</i>	<i>isbj</i>	<i>šepeš</i>	<i>espéjæf, f.</i>	<i>pil(o)</i>
mince-PR	<i>de-rijen</i>	<i>ar-injin</i>		<i>ænjin</i>	<i>enji (P^T)*</i>	<i>gæt</i>
sell-PR	<i>foruš</i>	<i>xruš</i>	<i>fruš</i>	<i>ruš</i>	<i>(h)ruš</i>	<i>osbj</i>
sell-PT	<i>foruxt</i>	<i>xrut</i>	<i>fruxt</i>	<i>rut</i>	<i>-yrut</i>	<i>esbij</i>
sickle	<i>daz</i>			<i>dáre, f.</i>	<i>dere</i>	<i>dære</i>

⁶ Gāzārxāni material is from Ivanow 1931.

	GILAKI	C. TATI	RUDBĀRI (TATOID)	S. TATI	TĀLÉQĀNI (TATOID)	<i>Alamut area⁶</i>	TĀLÉQĀN	C. CASP.
LEXEME						Dikin-Pāšeī	Gāzaxān	Gurāni
Rashī	Kaftejī-(Ketāsi)	Tutkāboni	Jubani	Chāli				Kelārdāšti
son-in-law	<i>damad</i>	<i>zama</i>	<i>zama</i>	<i>zo:ma</i>	<i>zama</i>	<i>zæma</i>	<i>zeyma</i>	<i>zoma</i>
sparrow	<i>čičini</i>	<i>mærgijə</i>	<i>mæljię</i>	<i>veskénjae</i>	<i>čuček</i>	<i>melijęe</i>	<i>čičok</i>	<i>miška</i>
wolf	<i>gurg</i>	<i>wærg</i>	<i>gorg</i>	<i>vérgeę, f.</i>	<i>verg</i>	<i>werk</i>	<i>verg</i>	
Alternates	*Gâleši	*Mâsali		*Tâkestâni	*Owrâzâni			

3. Phonology

3.1. Consonants

The general consonant inventory for most of the languages presented here is:

Table 3A

	Labial	Alveolar	Palatal	Velar	Uvular	Dental/ Laryngeal
voiceless	p	t	č	k		(?)
voiced	b	d	j	g	q [g]	
voiceless	f	s	š		x [χ]	
voiced	v	z				h [h]
	m	n				
	r					
	l					
	y					

General observations:

- In all the languages of these groups, voiceless stops, except for /?,/ are aspirated.
- In most languages /k/ and /g/ are palatalized before front vowels and word-finally. In the Vafsi of Vafs, velars are not heavily palatalized but they are so in Gurchān Vafsi. In Gilaki, most of Talyshi, and a few neighboring Tati languages, these consonants are always unpalatalized. Palatalization has gone so far in Sagzābādi that most velars have merged with the palatals before front as well as back vowels, but not consistently: [Front] (voiceless) čærð ‘did’, mæmlečæt ‘country’, čöštæ ‘killed’, mömčen ‘possible’; (voiced) u-jir ‘pick up!’, vaeljæ ‘leaf’; [Back] (voiceless) čar ‘work’, ču(h) ‘mountain’, šečar ‘hunt(ing)’, čuæst ‘pounded’, (voiced) jow ‘want’ (< *gæv-); [Unpalatalized] veškenj ‘sparrow’, mekašest ‘pulled’, væk ‘kidney’; čærgæ ‘hen’, a-gærdæst ‘he returned’, gow ‘cow’ (see also Stilo 1994).

In Harzani, as in the local Azeri dialect, the palatalization of /k/ in final position yields a subphonemic mid-palatal voiceless fricative [ç]; /čök/ [čöç] ‘good’. Sokolova (1953: 113–114) lists non-palatalized vs. palatalized velars (/k/ vs. /č/, /g/ vs. /g/) as separate phonemes in N. Talyshi. Neither Miller nor Pirejko list them separately and there is no accommodation for them in the Latin or Cyrillic alphabets used for Talyshi. In my own observations—although I never singled this issue out for investigation—I have found that palatalized velars exist independently in place names and in Azeri loans that are only partially integrated into the Talyshi lexicon. Hence these phones should probably be classified as marginal phonemes with a very low functional yield of the palatalized variants.

- In Mazanderani and the languages of the plateau, the uvular /q/ (always voiced) has 1) a stop phone [q] in initial position or when geminated, 2) a fricative phone [χ] (henceforth γ) in intervocalic position, and 3) either variant in final position. I know of no contrasting examples of [q] vs. [γ] in the languages of this area.
In Gilaki and Talyshi, except in cases of heavy bilingualism with Persian, there is no uvular stop variant and the phoneme only has the variant /γ/ throughout: Rashti: *yurbán* [yurbé[>]n]⁷ ‘sacrifice’; Lāhijāni: *yolám* [yelém] ‘pen’; Māsulei: *yævi* ‘strong’. This feature is a hallmark of a Gilaki or Talyshi accent in Persian.
- My observation is that /h/ is voiced [ɦ] throughout the area.
- The glottal stop is highly marginal and rare. It occurs in words of Arabic origin and is a feature of erudite (and bilingual) speech and is generally elided.
- When an original syllable-final /h/ or /ʔ/ is elided, compensatory lengthening in the neighboring vowel occurs: Sāravi řæ:r (~ řæhr) ‘city’, mæ:n (~ mæn?) ‘obstacle’; Lāhijāni ma:kəmə (~ mahkəmə) ‘tribunal’; ta:rif (~ taʔrif) ‘definition’, Vafsi qæ:r (~ qæhr) ‘anger’, mæ:ni (~ mæʔni) ‘meaning’.

Additions and permutations to the phoneme inventory above:

- Some languages have a /ž/, contrasting with both /š/ and /j/, but it generally has a low functional load. Exx: Leriki žæ ‘hit (pl.)’ – jæv ‘barley’ – šæ ‘gone’, šæž ‘you have gone’ – bevæj ‘bad’ – šæš ‘six’; Asālemi že ‘to hit’ – jego ‘place’ – ſe ‘to go’; námæž ‘I don’t hit’ – námæš ‘I don’t go’ – ávæj ‘pull off!'; Koluri žen ‘lost’ – jel ‘cheek’ – ſet ‘milk’. There is no strong evidence for /ž/ in Mazanderani, Gilaki, or S. Tati. The four examples that Jahāngiri (2003: 28) adduces mostly occur in the context VžC and probably do not contrast with /j/ (further research is needed on its status). The Tonekāboni /ž/ phoneme (Amini 2003: 43–44) has an extremely low load functional load. Vafsi has no /ž/ phoneme although it does occur as an allophone of /j/.
- Vafsi has a voiceless fricative trill phoneme /ř/ with low functional yield. It consists of a simultaneous voiceless alveolar trill and voiceless alveolar-apical fricative, somewhat similar to the Czech phone /ř/ but voiceless. It occurs only medially, either VřV (*ařé* ‘mill’ *guřačenlē* ‘badger’) or CřV (*bínříč* ‘I bought’), although it perhaps also occurs initially.
- Vafsi /v/ has a frequent allophone [w] which (along with the glides *ew* and *əw*) is emblematic for Vafsi speakers. It occurs intervocally next to a back vowel, most prominently /a/. The other vowel may be a non-high front or back vowel. Its distribution is particularly obvious in morphological contexts, e. g., *ærwendom* ‘I find’ vs. *háwend* ‘find!’, *báve* [báwe] ‘s/he came’ (with historical (*m* > *v/w*). When the other vowel is /æ/ (*æva*, *avæ*), the whole sequence changes

⁷ The superscript symbols [>] and [<] denote retracted and advanced vowel qualities respectively.

to *awa*: *aervazom* > *báwazom* ‘I say (present > Subjunctive)’ < **bévazom*; *yæv* ‘barley’ vs. *yawá* ‘threshed barley’, but note the feminine morpheme in *áwæ* ‘water’. It seems, however, that the word *vávæ* ‘almond’ (**váwæ*) contrasts with *áwæ*, thus conferring phonemic status to both /v/ and /w/, but a fuller statement must remain for future investigation. When only front vowels are present, the [w] phone is not used: *bævæ* ‘take!’, *dív/e* ‘face/s’. The [v] allophone is also retained in sequences of -iva-, e.g., *ivan* ‘threshing floor’, but not -avi- > -ewi-.

3.2. Vowels and Glides

There is a general split in these languages between a 6-vowel and 7-vowel inventory, plus various diphthongs. In the plateau area, essentially most of Tati, there is often a strict division of front vs. back, with no central vowels. Some languages, however, have a central, a high back unrounded, or a front rounded vowel as a seventh vowel. An eighth phoneme /ü/ is quite rare, but occurs as a common allophone of /u/.

Table 3B

Vafsi, Hezārrudi Gandomābi, Nowkiāni	Razajerdi, Alviri, Chāli Dikin-Marāqe, Kajali	Koluri, Māsulei Kafteji-Kelāsi
[F B]	[F B]	[F C B]
i u	i u	i u
e o	e ö o	e ə o
æ a [v]	æ a [v]	æ a [v]

Rashti, Lahijāni	Mazanderani	Leriki
[F C B]	[F B]	[F B]
i u	i u	i ī u
e ə o	e o	e o
←a→	ε	←a→ ←a→
	æ a [v]	

One unclear issue for Mazanderani in general is the distinction of two front vowels /e/ and /ɛ/. The two main grammars of Sāravi, Shokri (1995) and Yoshie (1996), make no mention of this distinction, but it is made for other dialects in Rastorgueva and Edel'man (1982), Borjian and Borjian (2007b), Lambton (1938), and in my own field notes for Bābolsari.

Minimal (or near-minimal) pairs in various dialects: Kordkheyli (Borjian) *te* ‘you’ vs. *te* ‘your’; *me* ‘my’ vs. *me* ‘I (subject)’; *véne* ‘must, should’ vs. *véne* ‘his, her, its’; *aere* ‘yes’ vs. *kære* ‘butter’; Bābolsari (Stilo field notes) *te* ‘your(s)’ vs. *te* ‘you’; *se* ‘apple’ vs. *se* ‘three’; *me* ‘my’ vs. *men* ‘I’; *šemé* ‘your (pl.)’ vs. *jémé* ‘shirt’; *vésse* ‘for’ vs. *véssene* ‘it tears’; Velātrui (Lambton) *véni* ‘must’ vs. *vení*

'nose'; (Rastorgueva & Edel'man) *də* 'two' vs. *de(h)* 'village' vs. *dæ* 'ten'; *ser* 'sated' vs. *serx* 'red'. Shokri, however, does occasionally show a type of contrast: *se* 'three' vs. *se:* 'apple'; *ser* 'secret'/*serx* 'red' vs. *se:r* 'sated'. Contrasts are rare but a token count of a portion of Borjian's corpus shows the low functional load of /e/: ε 2228, a 971, e 730, æ 494.

In Leriki /æ/ can be as far front as in Persian or Azeri but both /æ/ and /a/ are mostly articulated slightly more centrally while still retaining their distinction: *da* 'ten' vs. *dæ* 'valley'; *ha* 'each' vs. *hæ* 'donkey'; *hani* 'sleep (oblq.)' vs. *hæni* 'still, yet'.

General observations on Gilaki vowels:

- /e/ and /o/ have a rather low functional load in Gilaki, slightly higher in Lāhijāni.
- An expected /ə/ changes to /a/ in contact with /h/ (which may then be elided in postvocalic position yielding compensatory lengthening: Lāhijāni *hasən* 'Hasan', *ma:kəmə* 'tribunal' (NJ:35)).
- As Figure 2 shows, the /a/ and /ə/ phonemes have a wide range of free-varying allophones; the phone [æ] is much more common in Lāhijāni than in Rashti. The allophones of /a/ and /ə/ are often difficult to distinguish, e. g., *hæsán=a* (Objective) and *hæsán=ə* (Possessive) often sounded the same.
- For interest's sake, the vowels of the Lāhijāni text in §10.1 are left in allophonic transcription.

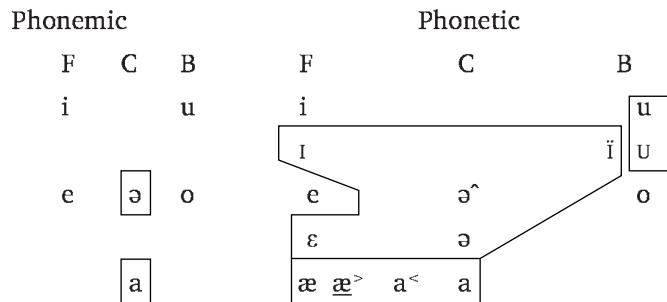


Figure 2: Phonetics and phonology of Lāhijāni vowels

Examples for Fig. 2:

- /ə/: [ə] *bəyn* ‘look!’, *békət̪i* ‘he fell’, *dávədənəm* ‘I close’, *sər ~ ser* ‘head’
 [ɪ] *óri* ‘there’, *šən̪i* ‘goes’, *búdm* ‘they did’, *xənd̪i* ‘laughter’; *dárt̪i* ‘is in’
 [ɛ] *yelém* ‘pen’, *búte ~ -ti* ‘he said’, *bá- ~ békhem* ‘I bought’, *mién* ‘(with) in’, *yézá* ‘food’, *iraj* [*irej*] ‘man’s name’
 [ɪ] *rík̪i* ‘boy’, *bí~bé~bámæ* ‘he came’, *bišun* ‘(that) he go’,
ríng~rēng~rəng ‘color’
- /a/: [a] *béma~bámæ* ‘he came’, *zakón* ‘children’, *dubaré* ‘again’, *zæmán* ‘time’
 [æ] *tifáeng* ‘gun’, *æmó* ‘we’, *nækun* ‘don’t do!’, *áegə ~ áege* ‘if’, *áeni* ‘you come’, *dóvæss’em* ‘I closed’, *haesán* ‘P.N.’, *mæ* ‘me (oblg.)’
 [a[˘]] *isæm~isa[˘]m* ‘I am at (Be₃)’
- /u/: [U]: *búšo* ‘s/he went’, *búgom* ‘(that) I say’, *turš* ‘sour’, *púrsi* ‘you used to ask’
 [u]: *búte* ‘s/he said’, *mu* ‘I’, *yuri* ‘teapot’, *ruz* ‘day’, *pul* ‘money’, *bum* ‘I was’

Glides and other additions and permutations in the above vocalic systems:

Vafsi: The Vafsi mid, front vowel /e/ is realized phonetically as a higher [I] in all positions, except with off-glides. Vafsi glides are: /ey ~ æy/, the marginal /uy/, /oy/, /ay/, and the rather distinctive (and emblematic for the Vafsi speakers) /əw/ and /ɛw/. The frequencies of each Vafsi vowel (in tokens) from a small corpus containing 6579 vowel nuclei are as follows: /æ/ 1830; /e/ 1595; /a/ 1244; /i/ 1138; /o/ 580; /u/ 192.

Vafsi vowels and glides:

- | | | | |
|----|--|-----|---|
| æ: | <i>ærvaæzóm</i> ‘I jump’, <i>ayré</i> ‘hives’ | ey: | <i>eyd</i> ‘holiday’, <i>leyléy</i> ‘boy (oblg.)’ |
| e: | <i>esdaré</i> ‘star’, <i>qeréq</i> ‘frost’ | uy: | <i>šuy</i> ‘husband (oblg.)’ (rare and avoided) |
| a: | <i>ána</i> ‘that direction’, <i>ban</i> ‘roof’ | oy: | (rare) <i>roy</i> ‘place name’ |
| i: | <i>in</i> ‘this’, <i>vilí</i> ‘sleet’ | ay: | <i>ayr</i> ‘fire’, <i>vaym</i> ‘I said’, <i>kay</i> ‘where’ |
| o: | <i>orián</i> ‘mumps’, <i>goló</i> ‘throat’ | əw: | <i>əwr</i> ‘cloud’, <i>vəwsı</i> ‘Vafsi’, <i>təw</i> ‘taste’ |
| u: | <i>uí</i> ‘hope’, <i>duriú</i> ‘lie’ | ew: | <i>ewí</i> ‘open field’, <i>hewlí</i> ‘donkey foal’, <i>hæréw</i> ‘hæram, ritually unclean’ |

Vafsi consonants:

p:	<i>pilæ</i> ‘pod’, <i>hápærs</i> ‘ask!’, <i>čærp</i> ‘fatty’	s:	<i>seza</i> ‘voice’, <i>ræsænæ</i> ‘rope’, <i>luas</i> ‘fox’
b:	<i>bæbékæ</i> ‘pupil (eye)’, <i>qærb</i> ‘grave’	z:	<i>zellé</i> ‘woman’, <i>azan</i> ‘this way’, <i>æz</i> ‘I’
t:	<i>tútæ</i> ‘mulberry (f.)’, <i>kirfit</i> ‘match(es)’	š:	<i>šišæk</i> ‘2-year-old lamb’, <i>qeyš</i> ‘belt’
d:	<i>dade</i> ‘older sister’, <i>gæsd</i> ‘bad’	x:	<i>xerxelle</i> ‘esophagus’, <i>feráx</i> ‘wide’
č:	<i>či</i> ‘fodder’, <i>geræč</i> ‘chalk’	h:	<i>hólæ</i> ‘ash’, <i>ahén</i> ‘iron’, <i>dæh</i> ‘ten’
j:	<i>jujíæ</i> ‘chick’, <i>qolenj</i> ‘cramp	m:	<i>mama</i> ‘midwife’, <i>čæmm</i> ‘eye’
k:	<i>kokk</i> ‘fat’, <i>tik</i> ‘thorn’	n:	<i>nošoáni</i> ‘evening’, <i>ætén</i> ‘now’
g:	<i>gælðégæ</i> ‘basket’, <i>værg</i> ‘wolf’	r:	<i>rékkæ</i> ‘line’, <i>širišæ</i> ‘leek’, <i>ayr</i> ‘fire’
q:	<i>qereq</i> ‘frost’, <i>aqáte</i> [ayáte] ‘speech’	ř:	<i>ařé</i> ‘mill’
f:	<i>ſérfæ</i> ‘snow’, <i>saf</i> ‘smooth’	l:	<i>leylæ</i> ‘boy’, <i>bol</i> ‘tarantula’
v:	<i>vávæ</i> ‘almond’, <i>gændev</i> ‘wheat’	y:	<i>yúæ</i> ‘heifer’, <i>xeyar</i> ‘cucumber’

Māsulei vowels and glides:

æ:	<i>æmæ</i> ‘we’, <i>šæm</i> ‘we go’	u:	<i>uškám</i> ‘I split open’, <i>pučú</i> ‘cat’
e:	<i>emé</i> ‘this (obl.)’, <i>xelék</i> ‘spade’	ey:	<i>veyv</i> ‘bride’, <i>šey</i> ‘shirt’
ə:	<i>əštá</i> ‘your’, <i>baz</i> ‘goat’	æy:	<i>æy</i> ‘him (oblq.)’, <i>pæydo</i> ‘apparent’
a:	<i>adoa</i> ‘had given’, <i>xam</i> ‘I want’	uy:	<i>uy</i> ‘he would come’, <i>xuyš</i> ‘you wanted’
i:	<i>illæ</i> ‘one, a’, <i>vieri</i> ‘he would take’	oy:	<i>paltóy</i> ~ <i>paltói</i> ‘a coat’
o:	<i>ošan</i> ‘tonight’, <i>zonó</i> ‘knee’	ay:	<i>aynæ</i> ‘mirror’, <i>xay</i> ‘you want’

Māsulei consonants:

p:	<i>pierám</i> ‘I pick up’, <i>veprazəm</i> ‘I lean’	z:	<i>zonám</i> ‘I know’, <i>væzám</i> ‘I jump’, <i>æz</i> ‘I’
b:	<i>bærá</i> ‘brother’, <i>lablab</i> ‘barking’	š:	<i>šen</i> ‘to go’, <i>xorošám</i> ‘I sell’, <i>bánəš</i> ‘sit!’
t:	<i>taxt</i> ‘galloping’, <i>šátəmæ</i> ‘I was able’	ž:	<i>žæná</i> ‘hits’, <i>bážænəm</i> ‘(that) I hit’
d:	<i>darám</i> ‘I have’, <i>šædíd</i> ‘intense’	x:	<i>xər</i> ‘fog’, <i>dáxon</i> ‘call!’, <i>báčærx</i> ‘turn!’
č:	<i>čem</i> ‘eye’, <i>əčára</i> ‘why?’, <i>hič</i> ‘none’	γ:	<i>γaz</i> ‘goose’, <i>dæyóštəmæ</i> ‘I drew water’,

j:	<i>jevæjám</i> ‘I pull down’, <i>ávæj</i> ‘pull up’	ræsey	‘friend’
k:	<i>katek</i> ‘cottage’, <i>čæká</i> ‘it breaks’	h:	<i>hærám</i> ‘I eat’, <i>gahen</i> ‘cows’
g:	<i>gusénd</i> ‘sheep’, <i>agærdám</i> ‘I return’, <i>xesárg</i> ‘mother-in-law’	m:	<i>máemon</i> ‘don’t stay’, <i>am</i> ‘I come’
f:	<i>færš</i> ‘carpet’, <i>kæfá</i> ‘up’, <i>sef</i> ‘apple’	n:	<i>næm</i> ‘I put’, <i>venen</i> ‘they see’
v:	<i>vervoærəm</i> ‘I unload’, <i>lev</i> ‘leaf’	r:	<i>ruk</i> ‘small’, <i>bæroær</i> ‘brother (obl)’
s:	<i>siá</i> ‘it burns’, <i>æsárg</i> ‘tear’, <i>péles</i> ‘lick!’	l:	<i>lækám</i> ‘I fall’, <i>kilil</i> ‘key’
		y:	<i>yen</i> ‘wife’, <i>dayen</i> ‘aunt (MoBroWi)’

Lāhijāni vowels and glides:

ə:	<i>dávæd</i> ‘close!’, <i>riká</i> ‘boy’	əy:	<i>bæys</i> ‘stay!’, <i>alnáy</i> ‘right now’
e:	<i>ére</i> ‘here’, <i>čére</i> ‘why?’	ay:	<i>háytəm</i> ‘I got’, <i>kočtay</i> ‘child’
a:	<i>ánəm</i> ‘I come’, <i>háda</i> ‘s/he gave’	oy:	<i>lakóy</i> ‘girl’
i:	<i>ísi</i> ‘you are’, <i>tínəm</i> ‘I can’	uy:	<i>kuyta</i> ~ <i>koyta</i> ‘which one?’
o:	<i>óre</i> ‘there’, <i>xoná</i> ‘house’, <i>æmo</i> ‘we’	aw:	<i>aw</i> ‘water’
u:	<i>úsan</i> ‘pick up!’, <i>murjú</i> ‘lentil’	ow:	<i>ištóws</i> ‘he heard’, <i>gow</i> ‘cow’

Lāhijāni consonants:

p:	<i>pildaná</i> ‘big’, <i>bápurs</i> ‘ask!'	z:	<i>zak</i> ‘child’, <i>bázən</i> ‘hit!’, <i>ayúz</i> ‘walnut’
b:	<i>bábin</i> ‘cut!’, <i>gəb</i> ‘talk’	š:	<i>šəmə</i> ‘you’, <i>ušón</i> ‘they’, <i>piš</i> ‘ago’
t:	<i>tódi</i> ‘throw!’, <i>bútəm</i> ‘I said’, <i>kot</i> ‘coat’	x:	<i>xaxor</i> ‘sister’, <i>surx</i> ‘red’
d:	<i>dávæd</i> ‘close!’, <i>hádam</i> ‘I gave’	γ:	<i>yelém</i> ‘pen’, <i>bayón</i> ‘gardens’, <i>duróy</i> ‘lie’
č:	<i>čére</i> ‘why?’, <i>báčin</i> ‘pick!’, <i>hič</i> ‘none’	h:	<i>hádi</i> ‘give!’, <i>báhem</i> ‘I bought’
j:	<i>jujə</i> ‘chick’, <i>xoj</i> ‘(wild) pear’	m:	<i>mi</i> ‘my’, <i>bómam</i> ‘I came
k:	<i>kərk</i> ‘chicken’, <i>laku</i> ‘girl’	n:	<i>nánəm</i> ‘I don’t come’, <i>livón</i> ‘glass’
g:	<i>gúnəm</i> ‘I say’, <i>ágə</i> ‘if’, <i>sang</i> ‘stone’	r:	<i>riká</i> ‘boy’, <i>óre</i> ‘there’, <i>pir</i> ‘old’
f:	<i>fúxus</i> ‘hide!’, <i>tiféng</i> ‘gun’, <i>tif</i> ‘thorn’	l:	<i>lakú</i> ‘girl’, <i>pəlá</i> ‘pilaff’, <i>jəngál</i> ‘forest’
v:	<i>víris</i> ‘get up!’, <i>dávæd</i> ‘tie!’, <i>div</i> ‘demon’	y:	<i>yad</i> ‘memory’, <i>diyát</i> ‘village(s)’
s:	<i>sib</i> ‘apple’, <i>kəsif</i> ‘dirty’, <i>durus</i> ‘right’		

Sāravi vowels and glides (Shokri 1995):

æ:	<i>aeger</i> ‘if’, <i>ræsen</i> ‘rope’, <i>dæ</i> ‘ten’	u:	<i>u</i> ‘water’, <i>lulu</i> ‘bogey man’
e:	<i>esbe</i> ‘white’, <i>kerk</i> ‘hen’	æy:	<i>æyd</i> ‘holiday’, <i>bæyr</i> ‘take!’
a:	<i>azad</i> ‘free’, <i>esa</i> ‘now’	ay:	<i>ayne</i> ‘mirror’
i:	<i>izi</i> ‘type of aphid’, <i>tim</i> ‘seed’	ow:	<i>dowr</i> ‘around’
o:	<i>oresi</i> ‘shoe(s)’, <i>kol</i> ‘dull’		

Sāravi consonants (Shokri 1995):

p:	<i>pærčim</i> ‘fence’, <i>hápers</i> ‘ask!’, <i>čæp</i> ‘left’	s:	<i>saze</i> ‘broom’, <i>lesek</i> ‘snail’, <i>kindes</i> ‘medlar’
b:	<i>bækub</i> ‘pound!’, <i>zebún</i> ‘tongue’	z:	<i>zezem</i> ‘wasp’, <i>kuze</i> ‘jug’, <i>maz</i> ‘bee’
t:	<i>tæli</i> ‘thorn’, <i>deter</i> ‘girl’, <i>gæt</i> ‘big’	š:	<i>šal</i> ‘jackal’, <i>vešna</i> ‘hungry’, <i>hénis</i> ‘sit!’
d:	<i>dægerd</i> ‘return!’, <i>badem</i> ‘almond’	x:	<i>xalik</i> ‘saliva’, <i>tælxun</i> ‘tarragon’, <i>sulax</i> ‘hole’
č:	<i>čečem</i> ‘type of weed’, <i>nuč</i> ‘sticky’	h:	<i>hæli</i> ‘plum’, <i>ahén</i> ‘iron’
j:	<i>jeme</i> ‘shirt’, <i>keje</i> ‘where’, <i>kelaj</i> ‘crow’	m:	<i>mi</i> ‘hair’, <i>nemašun</i> ‘evening’, <i>dim</i> ‘face’
k:	<i>kæk</i> ‘flea’, <i>esseka</i> ‘bone’, <i>fenizek</i> ‘mucus’	n:	<i>nemek</i> ‘salt’, <i>kenes</i> ‘stingy’, <i>kelen</i> ‘ash’
g:	<i>gæt</i> ‘large’, <i>aeger</i> ‘if’, <i>verg</i> ‘wolf’	r:	<i>rika</i> ‘boy’, <i>pe:ra</i> ‘day after tomorrow’, <i>mar</i> ‘mother’
q:	<i>qotele</i> ‘teeny’, <i>mæqez</i> [γ] ‘fly’, <i>daq</i> ‘hot’	l:	<i>læmbik</i> ‘termite’, <i>geles</i> ‘drool’, <i>gæl</i> ‘mouse’
f:	<i>fiye</i> ‘shovel’, <i>æšnoſe</i> ‘sneeze’, <i>jif</i> ‘pocket’	y:	<i>yæx</i> ‘ice’, <i>saye</i> ‘shade’
v:	<i>væk</i> ‘frog’, <i>kævez</i> ‘turtle’		

Leriki vowels and glides:

æ:	<i>ævæsór</i> ‘spring (season)’, <i>æmá</i> ‘we’	æy:	<i>æyšte</i> ‘to get up’, <i>pæyné</i> ‘day before yesterday’, <i>šæy</i> ‘shirt’
e:	<i>éfešin</i> ‘sneeze!’, <i>hæyé</i> ‘to itch’	uy:	(rare > u) <i>ruy</i> ‘river (oblq.)’
a:	<i>asp</i> ‘horse’, <i>bala</i> ‘cub’	oy:	<i>boylu</i> ‘little brother’, <i>moy</i> ‘fish’
i:	<i>ižnæn</i> ‘again’, <i>ævindi</i> ‘he would see’	ay:	<i>ayb</i> ‘shame’, <i>ovayná</i> ‘mirror’
í:	<i>ištæn</i> ‘self’, <i>angiští</i> ‘ring’	íy:	<i>giy</i> ‘neck’
o:	<i>oyó</i> ‘awake’, <i>astová</i> ‘star’	iy:	<i>žiy</i> ‘down’
u:	<i>umi</i> ‘hope’, <i>tul</i> ‘mud’, <i>lu</i> [lü] ‘peel’		

Leriki consonants

p:	<i>piyo</i> ‘over’, <i>bæpe</i> ‘must’, <i>bip</i> ‘quince’	z:	<i>zíne</i> ‘yesterday’, <i>væzæx</i> ‘frog’, az ‘I’
b:	<i>barz</i> ‘high’, <i>sibíž</i> ‘louse’, <i>limb</i> ‘snot’	š:	<i>šindi</i> ‘ladder’, <i>éfešin</i> ‘sneeze!’, <i>meš</i> ‘bee’
t:	<i>tæværs</i> ‘hail’, <i>kütlæ</i> ‘puppy’, <i>lipút</i> ‘lip’	ž:	<i>židé</i> ‘rope’, <i>bæží</i> ‘alive’, <i>nimáž</i> ‘noon’
d:	<i>di</i> ‘late’, <i>pédür</i> ‘rip!’, <i>péšand</i> ‘pour!’	x:	<i>xun</i> ‘blood’, <i>oxo</i> ‘end’, <i>væzæx</i> ‘frog’
č:	<i>čóknæ</i> ‘how?’, <i>nečí</i> ‘wolf’, <i>hič</i> ‘none’	γ:	<i>γandé</i> ‘to send’, <i>æγíl</i> ‘child’, <i>síy</i> ‘stone’
j:	<i>jæv</i> ‘barley’, <i>penjo</i> ‘fifty’, <i>bevæj</i> ‘bad’	h::	<i>havz</i> ‘green’, <i>éhašt</i> ‘hang!'
k:	<i>kæ</i> ‘house’, <i>ókæ</i> ‘open!’, <i>ak</i> ‘sweat’	m:	<i>mæsé</i> ‘to hear’, <i>æmæ</i> ‘we’, <i>yem</i> ‘fodder’
g:	<i>gæv</i> ‘mouth’, <i>égín</i> ‘fall!’, <i>kag</i> ‘hen’	n:	<i>nečí</i> ‘wolf’, <i>zoní</i> ‘knee’, <i>bíndín</i> ‘stand!'
f:	<i>fik</i> ‘thought’, <i>gofe</i> ‘cradle’, <i>lef</i> ‘quilt’	r:	<i>rívós</i> ‘fox’, <i>moræ</i> ‘mouse’, <i>bísř</i> ‘laugh!'
v:	<i>voš</i> ‘rain’, <i>ævon</i> ‘they’, <i>ov</i> ‘water’	l:	<i>kælæ</i> ‘bull’, <i>bíliv</i> ‘move!’, <i>hil</i> ‘hole’
s:	<i>sük</i> ‘rooster’, <i>osín</i> ‘iron’, <i>ars</i> ‘tear’	y:	<i>yol</i> ‘big’, <i>væyæ</i> ‘wedding’

3.3. Stress

When no inflectional morphology is involved, including feminine forms, lexical stress is word-final in most word classes: nouns, pronouns, adjectives, numerals, and non-finite verb forms, e. g., infinitives and participles. Conjunctions (e. g., *ége* ‘if’) and sometimes adverbs (e. g., *bélke* ‘perhaps’) allow for final or non-final stress.

Table 3C: Stress, main categories

	Noun	Numeral	Pronoun	Adjectives	Adverb
Sāravi	<i>riká</i>	<i>yazzé</i>	<i>emá</i>	<i>vešná</i>	<i>ferdá</i>
Lāhijāni	<i>riká</i>		<i>amó</i>	<i>gursəná</i>	<i>fərdé</i>
Rashti	<i>pəsár</i>	<i>yazdá</i>	<i>amá</i>	<i>vištá</i>	<i>pilé, pil(l)á</i>
Māsulei	<i>zoé</i>	<i>yunzé</i> (Māsāli)		<i>ænəštá/višír</i>	<i>sæbá</i>
Asālemi	<i>zué</i>		<i>æmé</i>	<i>veší</i>	<i>sævsér</i>
Leriki	<i>zoé</i>	<i>yonzé</i>	<i>æmé</i>	<i>væší</i>	mášti
Kalāsuri	<i>zuré</i>		<i>æmén</i>	<i>væší</i>	<i>yæké</i>
Koluri	<i>kulé</i>	<i>yanzé</i>	<i>æmé</i>	<i>višír</i>	<i>pillé</i>
Kelāsi	<i>koté</i>	<i>yazzé</i>	<i>amá</i>	<i>xará</i>	<i>pillé</i>
Chāli	<i>puré</i>	<i>yazzé</i>	<i>æmá</i>	<i>göšná</i>	<i>pil(l)é</i>
Vafsi	<i>leylé</i>	<i>yazdá</i>	<i>awán</i>	<i>gosná</i>	<i>mæzén</i>
	‘boy’	‘eleven’	‘we’	‘hungry’	‘large’
					‘tomorrow’

Languages with feminine gender in nouns (and in some languages, adjectives modifying them) place the stress routinely on the penultimate. This does not include the Vafsi Feminine II type (§5.2.1), which takes final-stress in the Direct singular.

Table 3D: Stress and nominal morphology

Vafsi	‘wheat stalk (F.DIR.SG.)’	<i>suálæ</i>	<i>sualé</i>	‘unglazed pot shard’ (M.OBL.SG.)
	‘hens (DIR.PL.)’	<i>kárgæ</i>	<i>kærgé</i>	‘hen (F.OBL.SG.)’
	‘big (F.DIR.SG.)’	<i>mæzænæ</i>	<i>mæzænéd</i>	‘the big one (M.DIR.SG.)’
	‘starlings (DIR.PL.)’	<i>čillé</i>	<i>čillé</i>	‘scarlet fever’ (FEMII.DIR.SG.)’
Kafteji	‘hens (DIR.PL.)’	<i>kárgø</i>	<i>kærgá</i>	‘hen (F.OBL.SG.)’

Since the punctual marker *bV-* often occurs in both the Preterit and the Past Participle, the former (a finite form) is stressed differently from the latter (a non-finite form). In Vafsi the stress pattern in the Present Perfect (see §6.5.9) generally contrasts with the feminine form of the Past Participle. In addition, the stress patterns for the Imperfect and the Past Subjunctive are different in Leriki Northern Talyshi.

Table 3E: Stress and verbal morphology

Vafsi	'sewn (PPL)' 'sewn (PPL, F)'	<i>bædutté</i>	<i>bééduttæ</i>	'X sewed' 'X has sewn'
Sāravi	'broken (PPL)'	<i>beškessé</i>	<i>béškesse</i>	'X broke' (SH: 40)
Asālemi	'risen (PPL)'	<i>æišté</i>	<i>æištæ</i>	'X has risen'
Leriki	'I picked up (IMF)'	<i>peægætím</i>	<i>peégætim</i>	'I picked up (PT.SBJ)'

Table 3F: Other random stress contrasts

Lāhijāni	'lady'	<i>xanóm</i>	<i>xánəm</i>	'I want'
Kordkheyli	'for'	<i>vessé</i>	<i>vésse</i>	'had to'
Leriki	'corn, maize'	<i>mæké</i>	<i>mákæ</i>	'don't do!'
Bābolsari	'type of insect'	<i>bæú</i>	<i>béú</i>	'become!'
Asālemi	'don't get up!'	<i>pe-máe-b</i>	<i>pé-m-æ-b</i>	'I get up'

3.4. Syllable structure

The common syllable structure throughout the area is (C)V(C)(C). Initial consonant clusters, as a general rule, are not allowed. Examples:

- V: Sāravi *u* 'water', *ema* 'we'; Lāhijāni *u* 'he', *óre* 'here'; Leriki *e* 'strawberry', *isæ* 'now'; Vafsi *awan* 'we', *éna* 'this way, in this direction'
- CV: Sāravi *še* 'self', *saze* 'broom', *lue* 'bark (dog)'; Lāhijāni *su* 'three', *rikə* 'boy'; Leriki *lo* 'crow', *væyu* 'bride', *zoæ* 'boy'; Vafsi *bi* 'again', *vávæ* 'almond', *luas* 'fox'
- VC: Sāravi *un* 'that'; Lāhijāni *alnáy* 'right now'; Leriki *ak* 'sweat'; Vafsi *æz* 'I'
- CVC: Sāravi; Lāhijāni *jir* 'down', *væsni* 'co-wife'; Leriki *gæv* 'mouth', *soynæ* 'shade'; Vafsi *bær* 'door', *kelje* 'girl'
- CVCC: Sāravi *šeft* 'crazy'; Lāhijāni *kərk* 'hen'; Leriki *merd* 'man'; Vafsi *gæsd* 'bad'

3.5. Gemination

While gemination is in principle possible across the whole area, it is often seen as a feature of erudite speech, esp. when occurring in Arabic loanwords. And indeed the gemination is usually retained in highly erudite words. Original geminated

consonants in most words are for the most part degeminated. There are significant exceptions.

Gemination in Sāravi tends to be less frequent than in other varieties: *hæsseka* ‘bone’. Geminated nasals created in the Present tense via morphology tend to undergo dissimilation in Sāravi but in other dialects are retained, e. g., Bābolsari *dun-n-i* = Sāravi *dundi* ‘you know’, **kæfemne* > Bābolsari *kæfemme* = Sāravi *kæfembe* ‘I fall’.

Gemination seems to be more frequent in some forms of Tati. It is fairly frequent in Vafsi: *koppel* ‘tiny’, *hottidéni* ‘bedtime’, *rékkæ* ‘line’, *méssæ* ‘fly’, *esbezze* ‘spleen’, *ušša* ‘pimple’, *æjjene* ‘s/he hits’, *ænnie* ‘s/he puts’, *holle* ‘hole’, *zarru* ‘child’ and final gemination is usually only retained when followed by another vowel in a suffix, enclitic, or even another word: *rott* ‘place’, *čæmm/čémme* ‘eye/s’, *verr* ‘crazy’. Koluri has *mæččæ* ‘lip’, *killá* ‘girl’, *assen* ‘anvil’, *fett* ‘full’. Chāli (Yarshater 1969: 55) has *béttæj* ‘run!’, *xæssé* ‘bone’, *xöyyé* ‘wooden shovel’. Dikin Marāqei has *killégæ* ‘girl’, *ænnán* ‘I put’, *néggion* ‘I am not’, *zollég* ‘boy’. Kelāsi has *bámmæ* ‘s/he died’, *bózzæ* ‘hit’, *pillæ* ‘big’ but very few other examples appear in my field notes. Gemination is not as common in Leriki as in other languages but seems to be retained in certain words where it occurs: *bællædo* ‘oak’. Note also *mækæ* ‘corn, maize’ vs. *mækkæ* ‘Mecca’.

3.6. Intonation

Due to space limitations, I refer the reader to the claim made in Stilo (2004b: 274–277) that languages of western Iran have five basic intonational contours requiring the recognition of five levels of pitch. Four of these pitch levels are contrastive on the clause level only and come only following the obligatory highest Level 5 pitch (usually the only one per clause), which is borne by the word in focus according to the information structure of the clause. In §8.2, points 4 and 5, I discuss two highly distinctive contours used for 1) sentence-initial adverbial clauses and relative clauses, which I mark with (,,) in my transcription system, and 2) right-shifted adverbial clauses. The latter forms a (less common) sixth contrastive intonational contour and a sixth (extra-low) pitch level.

4. Grammatical isoglosses

4.1. Grammatical isoglosses in the Caspian family

The Caspian family as a whole shows a systemic grammatical restructuring through convergence with Persian: loss of case (§5.1.1) and gender, the *=ra* Objective marker (§5.1.3), the merger of intransitive and transitive verb types (§6.1), Passivization (§6.10.2) among others. There are some areas, however, that have

resisted Persian influence, including important isoglosses that help define the Caspian group. Certain of these features are shared by some Talyshi-Tati-Tatoid dialects (#3–5) or most forms of Tatic (#1, 2):

1. Caspian languages generally retain left-branching word order: Adjective-HEAD (§5.5.1), Possessor-HEAD (§5.4.1), and Postpositions (§5.7). Due to recent heavy influence from Tehrani Persian, these word order domains are now in flux as well.
2. Dedicated preposed Possessive forms of pronouns (§5.1.2)
3. Set₂ possessive enclitics on nouns are lacking (see next Point)
4. No Set₂ (encoding A, P) in verb paradigms; as opposed to Tatic (for the most part), the Patient is never encoded in the verb.
5. Restructuring of verbal paradigms did not include the Persian TAM *me-/mi-*. By contrast, Caspian and Tatic languages have a wide range of TAMs (see §6.5.1).

Other isoglosses reflect divisions within Caspian. The base for comparison here is EG (Lāhijāni). Some of these isoglosses are shared areally between EG and other Caspian, Tati and/or the Semnan area group (1–9) or between WG and Tati (8–10):

Table 4: Isoglosses of Eastern Gilaki

Eastern Gilaki	As opposed to:	EG shares with:
1. -(V)n- Present tense marker	Ø marker, WG (§6.5.1)	rest of Caspian, Semnan
2. -an/-on plural marker	-ešan, Central Caspian	rest of Caspian
3. *-ft, -xt > -t(t)	retention of -ft, -xt, WG	most of Caspian, NWI
4. six ‘to be’ verbs (§6.8)	three ‘to be’ verbs (§6.8.1)	WG, (5 in S. Talyshi)
5. X-ka-Be ₄ -, Progressive II	missing, WG (CC?)	missing throughout?
6. preverb <i>ha-</i> ~ <i>hæ-</i>	<i>fa-</i> ~ <i>fæ-</i> , WG	most of Caspian, NWI
7. Be ₄ (‘be in’), Progressive I (§6.5.8.1)		all Caspian
8. Three main preverbs	Multiple Preverbs, WG	most of Caspian (§6.3)
9. Ø Imperfect marker (§6.5.7)	-i-, WG	most of Caspian
10. Stress on Preterit marker	stress on root, Preterit, WG	most of Caspian
11. -ena- conditional marker	missing, WG	Rāmsari, Tonekāboni
12. Distinct 2/3PL PAMs	Syncretism in Rashti	most NWI
13. Future: invariable ‘want’ + Subjunctive (§6.5.4)	Future: conjugated ‘want’ ‘want’+ Infinitive	missing in most NWI (exc. some CPD)
14. <i>mæn-ni-</i> negative of ‘can’	missing elsewhere	

4.2. Grammatical isoglosses in the Tatic family (not including Tatoid)

The Tatic family is highly conservative, as demonstrated by the list of features below, shared by most varieties. Many of these features are also found elsewhere in NWI.

Conservative Features in Tatic:

1. Grammatical gender (Stilo: forthcoming d), still found in half of the Tati group.
2. Masculine and Feminine singular forms (and occasionally a Feminine II) and the plural forms each have distinct Direct vs. Oblique cases (Stilo: 2008b).
3. A special Oblique formant *-ær* is used for kinship terms of either gender.
4. Oblique case is used to mark Object in the Present system fairly universally, except for Alviri (no case), Razajerdi, and Harzani (both with =*ra* marking).
5. Canonical Ergativity is quite robust in some languages: Oblique A, Direct P, and both coreferenced in verb. Others have vestigial or marginal Ergativity.
6. Distinct Intransitive and Transitive paradigms in tenses built on Past stem.
7. Set₂ (Oblique) PAMs co-index the Agent and Set₁ (Direct) PAMs often co-index the Patient, reflecting canonical nominal Ergativity in the Past (§7.3).
8. Regular retention of a synthetic passive formation (§6.10.2).
9. Most Tatic languages retain an Optative paradigm (§6.5.3).
10. Certain Tatic languages retain the 2nd singular Set₁ PAM *-iš* (cf. *-i* elsewhere).

Tatic Innovations:

11. The *Ezafe* may exist sporadically, seemingly a borrowing from Persian.
12. Postpositions are usual, but some languages have a mixed typology (§5.7.1).
13. Preposed Genitive case of pronouns and the interrogative *ki* ‘who’ formed with a prefix *čə-* ~ *əš-* (< **hača* ‘from’), quite regular across Tatic, (§5.2.2).
14. Set₂ possessive enclitics (§7.3) are highly restricted or are lacking in most of Talyshi and some neighboring Tati varieties (Taromi group, Nowkiani).
15. Present TAMs have wide variation, areally distributed (see §6.5.1)
16. Imperfect TAMs have wide variation, areally distributed (see §6.5.7)
17. Alignments in the Past range from Ergative to Nominative-Accusative, to complete loss of case (Neutral alignment).
18. Some languages have Double Oblique marking, in which both A and P are Oblique-marked, either as the sole alignment or as an alternate (§7.1.3.2).
19. Differential Object Marking (DOM): Oblique case-marking only for salient Ps.

4.2.1. Talyshi isoglosses

Features shared by all Talyshi dialects (most repeated from above list):

1. Two-term case system at least in the singular, also in plural in many dialects
2. Lack of gender
3. Special genitival cases in pronouns, formed with the prefix *či-* ~ *iš-*
4. 2nd singular *-iš* in at least some tense-aspect-modal paradigms
5. leftbranching NP adjuncts (GN, AN) as the dominant pattern
6. Imperfect suffix formant *-i-* (see also point 12 below)
7. Formal contrast between past system intransitive and transitive paradigms
8. Set₂ PAMs generally leftwardly mobile in the past

Features that distinguish some Talyshi dialects from others:

Diachronic Phonology

9. N. Talyshi loss of intervocalic and postvocalic *-r* (with exceptions): (Leriki) *šin* ‘sweet’, *kag* ‘hen’, *do* ‘tree’, but *pónæ* ‘last year’ vs. *pornæ* ‘calf’; (Leriki, Viznei) *vit-* ‘run’; *divo* ‘wall’, but (Leriki) *bæm-* vs. (Viznei) *bæræm-* ‘cry’.
10. Shift of low vowels: **a* (æ) > *ā* (a): (Leriki, Viznei) *kag* ‘hen’, *mard-* ‘die’, *asp* ‘horse’, *az* ‘I’, *gandəm* ‘wheat’; **ā* (a) > *o*: (Leriki, Viznei) *vot* ‘say’, *boy* ‘garden’, *osin* ‘iron’, *pónæ* ‘last year’, *sor* ‘year’, *ov* ‘water’, *do* ‘tree’, *divo* ‘wall’, *cok* ‘good’, *ro* ‘road’, *bozor* ‘market’.

Synchronic Morphology

11. Plural Direct/Oblique cases merged in N. Talyshi but distinguished in others.
12. Augment formant *æ-* for the Imperfect tense, common in all of Northern and Central Talyshi dialects but lacking in Southern Talyshi (and in Tati).
13. Leftward mobility of Set₁ PAMs in analytically-formed tenses (§7.4.3).
14. Many Present and Past verb stems merge in N. Talyshi, esp. those ending in *d/t* (Leriki: *bíhand* ‘read!’), but there is a small subset with very distinct forms, e. g., *žæn/žæ* ‘hit’, and a subset with an *i* formant (*čini*) ‘pick’ that ‘... includes all verbs bearing the causative suffixes *-n-* and *-ovn-*...’ (Kaye 2013: 201). Kaye also makes an insightful point on past *a* stems, never before directly addressed in the literature: the Preterit of *mand-é* (INF) is *mándim* ‘I stayed’, and *níšt-é* (inf) is *níštim* ‘I sat’, but the ‘1SG preterit of *rasé* is not **rásim*, but *rasáym* [‘I arrived’].’ (Kaye 2013: 205).

5. The noun and the noun phrase

5.1. Nominal morphology in Caspian

5.1.1. Case

The status of case in Caspian languages is a thorny issue. There are two opposing views, one supporting, and one disclaiming, the existence of formal cases. My own opinion favors the non-case analysis, especially based on points 5, 6, and 8.

The case for case

1. An Oblique (Windfuhr 1992: 29) or Genitive (Lecoq 1989: 304) case encodes possessor and adjectival adjuncts. In favor of this analysis is the interpretation of the unstressed *-ə ~ -e* as derived from an old Oblique case. As other support, Vafsi may also mark adjectival modifiers with an Oblique case (ex. 27).
2. Singular SAP pronouns in Lāhijāni have special Objective case forms, derived from the original *=(r)a* grammaticalized into a suffix *a* (see Table 5B).
3. SAP pronouns in Caspian, as well as 3SG for Mazanderani, have special possessive forms (see Table 5B) potentially classified as the Genitive case, indicating the possibility of a case system in Caspian languages.

The case for non-case

4. As the “Reverse *Ezafe*” (REZ) (Stilo 2001) is used with possessors, adjectives, and most postpositions, it looks more like a general linker than a case form. The *Ezafe* is head-marking and right-branching, the opposite of the Reverse *Ezafe*.
5. The Oblique plural *case* is usually fusional, i. e., Oblique and Plural in one, but in Caspian languages these categories are agglutinative: plural + REZ > *-on-ə*.
6. The most likely candidate for the origin of *-ə* is the *-æ* linker (§5.5.2), an areal feature: Asālemi *sər-ǽ ðésb* ‘red horse’, Koluri *šàl-æ dærzen* ‘shawl needle’, Cauc. Tat (SWI): *túrš-æ nu* ‘sour bread’ (Grjunberg 1963: 156), etc. WG mostly realizes an original *æ* as *-ə*: *mədrəsə* ‘school’. The Maz. REZ *-e* may also derive from *æ* as *e* often corresponds to the *æ*-vowel of other languages and the *ə* of Gilaki.
7. All forms of Gilaki also have the *Ezafe*, a linker of *postposed* adjuncts to a head noun (mostly in borrowings from Persian, but not exclusively). Likewise, the *Ezafe* clearly does not derive from a diachronic case form.
8. Since the *=raæ* enclitic attaches to the right-hand margin of the NP, in the right-branching *Ezafe* pattern it can be completely removed from the head noun.
9. Pronominal object marking is more similar to the enclitic *=ra* after nouns: WG: *má=ra* >EG *m-a*, WG *tá=ra* >EG *t-a*, Lāhij. *xoná* ‘house’ > *xon-á* (+ RA).
10. The Genitive of SAP pronouns may simply be seen as a possessive particle.

5.1.2. Caspian noun and pronoun morphology

Nouns in Caspian lack grammatical gender. Except for occasional Arabic plurals, there is one plural formant, derived from Caspian **an*. Yoshie (1996: 14) gives a regular plural formant -(*h*)*a*, probably recently borrowed from Persian, while the expected plural is “hardly employed” in Mazanderani. Shokri (1995) gives both types.

Table 5A: Gilaki and Mazanderani nominal morphology

Singular	Lāhijāni			Sāravi		
Base	<i>gul</i>	<i>zak</i>	<i>xoná</i>	<i>berar</i>	<i>damad</i>	<i>ketab</i>
Object.	<i>gúl=a</i>	<i>zák=a</i>	<i>xon-á</i>	<i>berár=(r)e</i>	<i>damad=(r)e</i>	<i>ketab=(r)e</i>
Possess.	<i>gúl-ə</i>	<i>zák-ə</i>	<i>xon-é</i>	<i>berár-e</i>	<i>damad-e</i>	<i>ketab-e</i>
Plural						
Base	<i>gul-ón</i>	<i>zak-ón</i>	<i>xon-án</i>	<i>berar-ún</i>	<i>damad-a</i>	<i>ketab-a</i>
Object.	<i>gul-ón=a</i>	<i>zak-ón=a</i>	<i>xon-án=a</i>	<i>berar-ún=(r)e</i>	<i>damad-a=re</i>	<i>ketab-a=re</i>
Possess.	<i>gul-ón-ə</i>	<i>zak-ón-ə</i>	<i>xon-án-ə</i>	<i>berar-ún-e</i>	<i>damad-a(-e)</i>	<i>ketab-a(-e)</i>
	‘flower’	‘child’	‘house’	‘brother’	‘son-in-law’	‘book’

In Central Caspian the plural formant is an unstressed -*ešan/ešon*: Rāmsari *dar/dárešon* ‘tree/s’, Tonekāboni *kija/kijášan* ‘girl/s’, Kelārdashti *rika/rikašon* ‘boy/s’. This is one of the major isoglosses establishing Central Caspian as a separate group.

Table 5B: Gilaki and Mazanderani pronominal morphology

Lāhijāni pronouns				Sāravi pronouns		
	Nominative	Objective	Possessive	Nominative	Objective	Possessive
1SG	<i>mu, mo</i>	<i>m-a</i>	<i>mi</i>	<i>men</i>	<i>me=re</i>	<i>me</i>
2SG	<i>tu</i>	<i>t-a</i>	<i>ti</i>	<i>tu</i>	<i>te=re</i>	<i>te</i>
3SG.PROX	<i>i(n)</i>	<i>ín=a</i>	<i>ín=ə</i>	<i>ve</i>	<i>ve=re</i>	<i>ve</i>
3SG.DIST	<i>u(n)</i>	<i>ún=a</i>	<i>ún=ə</i>			
1PL	<i>amə, amo</i>	<i>amə=ra</i>	<i>ame</i>	<i>ema</i>	<i>ema=re</i>	<i>æme</i>
2PL	<i>šəmə, šəmu</i>	<i>šəmə=ra</i>	<i>šime</i>	<i>šema</i>	<i>šema=re</i>	<i>šeme</i>
3PL.PROX	<i>išón</i>	<i>išón=a</i>	<i>išón=ə</i>	<i>vešún</i>	<i>vešún=re</i>	<i>vešún-e</i>
3PL.DIST	<i>ušón</i>	<i>ušón=a</i>	<i>ušón=ə</i>			

The Lāhijāni long forms (*in, un*) appear when a following suffix or enclitic begins with a vowel. The short forms (*i, u*) also function as demonstrative adjectives:

- (01) *tin-i i kar=a me=bə bú-kon-i? i ti pəsér=ə?*
can-2s this work-RA I=for PU-do-2s this your son=COP:3s
‘Can you do this thing for me?’ ‘Is this/he your son?’

In Mazanderani and most of Central Caspian we find the only languages in a vast area of Iran to have 3rd person pronouns that are not synchronic demonstratives (*ve/vešün*, etc.) These forms are used for both animates and inanimates. Kelārdaštī and Tonekābonī have both types, while Rāmsari is outside the isogloss and seems to have only synchronic demonstratives as 3rd person pronouns.

(02) Sāravi

<i>ve ſe zen-á=re kúš-en-e.</i>	<i>ve gusfén-e gušt=e.</i>
he own wife-PL=RA kill-PR-3S	it sheep-REZ meat=COP:3S
'He kills his wives.' (SY: 54)	'It is mutton.' (SY: 68)

5.1.3. The uses of the objective marker =ra

The enclitic =RA flags (specific) Direct Objects (see Alignment, and DOM, §7.1), Indirect Objects (§7.5), Experiencers (§6.7), and other Dative-like usages (§7.6).

Nouns/Third person pronouns

(03) Lāhijāni: Definite direct object

<i>mu diruz hæsán=a bá-de-m.</i>
I yesterday P.N.=RA PU-SEE:PT-1s
'I saw Hassan yesterday.'

Lāhijāni: Indirect Object

<i>mu hæsán=a pul há-da-m.</i>
I P.N.=RA money PVB-give:PT-1s
'I gave Hassan some money.'

(04) Sāravi: Definite Direct Object (SY: 68)

<i>šeter gušt=re koje=jem bá-it-i?</i>
camel meat-RA where=from PU-take:PT-2s
'Where did you get the camel meat from?'

Sāravi: Experiencer (SH: 149)

<i>aeslaš me=re særð næ-vu-n-e.</i>
at.all I=RA cold NEG-become-3s
'I absolutely never get cold.'

5.2. Nouns and pronouns in Tatic: case, gender, number

5.2.1. Noun morphology

As mentioned (§4.2), almost all Tati languages and dialects have retained a two-term (Direct/Oblique) nominal case system. In half of the Tati group, gender has also been retained in the singular, while it is lost in the rest of Tatic. The only languages to have lost the case system but still retain gender distinctions are Razajerdi and Alviri. Both Tatoid groups have lost all original traces of both case and gender (but see §5.4.4).

Table 5C: Tatic languages distinguishing gender

VAFSI			DIKIN MARĀQEI			KAFTEJI			CHĀLI	
Masc.	Fem. I	Fem. II	Masc.	Fem.	Masc.	Fem.	Masc.	Masc.	Fem.	
DIR.SG. <i>æsb-Ø</i>	<i>kérg-æ</i>	<i>kelj-é</i>	<i>xordé</i>	<i>kellég(-an)</i>	<i>mærdák-Ø</i>	<i>kelék-æ</i>	<i>bær-Ø</i>			
OB.SG. <i>æsb-i</i>	<i>kærg-é</i>	<i>kelj-i</i>	<i>xordég-i</i>	<i>kelleg-é</i>	<i>mærdák-ə</i>	<i>kelek-ə</i>	<i>bær-e</i>		<i>xʷák-æ</i>	
DIR.PL. <i>æsb-e</i>	<i>kérg-e</i>	<i>keljí-e</i>	<i>xordeg-in</i>	<i>kelleg-in</i>	<i>mærdák-ə</i>	<i>kelék-ə</i>	<i>bér-e</i>		<i>xʷák-e</i>	
OB.PL. <i>æsb-án</i>	<i>kærg-án</i>	<i>kelji-án</i>	<i>xordeg-ón</i>	<i>kelleg-ón</i>	<i>mærdæk-ón</i>	<i>kelek-ón</i>	<i>bær-ó(n)</i>		<i>xʷák-ó(n)</i>	
'horse'		'hen'		'girl'		'child'		'girl'		
7 distinct forms		6 distinct forms		5 distinct forms		4 forms				

Table 5D: Tatic languages not distinguishing gender

TATI				TALYSHI					
Khoini		Kelāsi		Kalāsuri	Leriki	Viznei	Asālemi	Māsulei	
DIR.SG. <i>pæs</i>	<i>mærdák-Ø</i>	<i>kelék-Ø</i>	<i>pæs</i>	<i>kag-Ø</i>	<i>kag-Ø</i>	<i>kærg-Ø</i>	<i>kærg-Ø</i>		
OB.SG. <i>pæs-é</i>	<i>mærdák-e</i>	<i>kelék-e</i>	<i>pæs-i</i>	<i>kag-i</i>	<i>kag-i</i>	<i>kærg-i</i>	<i>kærg-i</i>		
DIR.PL. <i>pæs-é</i>	<i>mærdæk-ón</i>	<i>kelek-ón</i>	<i>pæs-é</i>	<i>kag-ón</i>	<i>kag-é</i>	<i>kærg-é</i>	<i>kærg-én</i>		
OB.PL. <i>pæs-án</i>			<i>pæs-ún</i>		<i>kag-ón</i>	<i>kærg-ún</i>	<i>kærg-ón</i>		
'sheep'		'man'		'girl'		'sheep'		'hen'	
3 distinct forms each		4 forms		3 forms		4 forms		4 forms	

5.2.2. Pronoun paradigms

SAP pronouns and the interrogative *ki* 'who' have special pronominal strategies for Genitive formation (§4.2, Pt. 13). The Genitive forms of pronouns in Vafsi and some other Tati types have totally subsumed the original Oblique forms and their functions.

Table 5E: Vafsi SAP pronouns and ‘who’; Set₁ and Set₂

	Direct	Oblique	Set _{1A} (Direct)	Set _{1B} (Copula)	Set _{2A} (Oblique)	Set _{2B} (Prefixes)
1SG	<i>æz</i>	<i>tæ-men</i>	-om(e)	=im(e)	=om	<i>im-</i>
2SG	<i>tæ</i>	<i>es-dæ</i>	-i	=i	=i	<i>i-</i>
3SG.PROX	<i>in</i>	<i>t-in-i*</i>	-e ~ Ø	=e	=es	<i>is-</i>
3SG.DIST	<i>an</i>	<i>t-an-i*</i>				
1PL	<i>awan</i>	<i>ta-wan</i>	-am(e)	=am(e)	=oan	<i>oan-</i>
2PL	<i>soan</i>	<i>Ø-soan</i>	-a	=a	=ian	<i>ian-</i>
3PL.PROX	<i>ín-e</i>	<i>t-in-án*</i>	-énd(e)	=énd(e)	=esan	<i>isan-</i>
3PL.DIST	<i>án-e</i>	<i>t-an-án*</i>				
‘who’	<i>ki</i>	<i>t-e-g-é</i>				

*(3SG/PL Oblique alternates: *in-t-in-i*, *an-t-an-i*, *an-t-an-án*, *in-t-in-án*)

Table 5F: Leriki SAP pronouns and ‘who’; Set₁ and Set₂

				Set _{1A}	Set _{1B}	Set ₂
	Direct	Obliq.	Accus.	Genitive (Direct)	(Copula)	(Oblique)
1SG	<i>az</i>	<i>mí(n)</i>	<i>mí-ní</i>	<i>čí-mín</i>	-im	=im
2SG	<i>tí</i>	<i>tí</i>	<i>tí-ní</i>	<i>iš-tí</i>	-iš~iž	=e
3SG	<i>æv</i>	<i>æy</i>	(<i>æy-i?</i>)	<i>č-æy</i>	-i	=iš~iž
1PL	<i>æmæ</i>		<i>æmæ-ní</i>	<i>č-æmæ</i>	-imon	=imon
2PL	<i>šímæ</i>		<i>šímæ-ní</i>	<i>šímæ</i>	-išon~ižon	=on
3PL	<i>ævon</i>		<i>ævon-í</i>	<i>č-ævon</i>	-in	=išon~ižon
‘who’	<i>ki</i>		<i>ki</i>	<i>čí-ki</i>		

Leriki/Northern Talyshi (maximum: 4-term, 1SG)

Table 5G: Hezārrudi pronouns (maximum: 5-term, 1SG, 3PL)

	Subject	Pres/Obj	Past/Obj	Poss	Past/Ag.	Nº of forms
1SG	<i>æz</i>	<i>ædem</i>	<i>æčem</i>	<i>čem(e)(n)</i>	<i>men</i>	5
2SG	<i>tæ</i>	<i>ætæ</i>	<i>ætæ</i>	<i>eštæ</i>	<i>tæ</i>	3
3M	<i>o</i>	<i>æjo</i>	<i>o</i>	<i>jo</i>	<i>o</i>	3
3F	<i>oæ</i>	<i>æjo</i>	<i>oæ</i>	<i>jo</i>	<i>o</i>	3
PROX	<i>æm</i>	<i>æjéme</i>	<i>æm</i>	<i>jeme</i>	<i>æm</i>	3
3F	<i>démæ</i>	<i>æjemé</i>	<i>áémæ</i>	<i>jeme</i>	<i>démæ</i>	3
1PL	<i>æma</i>	<i>æčema</i>	<i>æčema</i>	<i>čema</i>	<i>æma</i>	3
2PL	<i>šema</i>	<i>æšema</i>	<i>æšema</i>	<i>šema</i>	<i>šema</i>	2
3PL	<i>o</i>	<i>æjuan~æjan</i>	<i>oe</i>	<i>juan-jan</i>	<i>oan</i>	5

Table 5H: Kafteji (maximum: 4-term, 1SG, 3SG)

	Subject:	Pres/Obj	Past/Obj	Poss	Past Agent
1SG	æz	čəmən	čəmən~čem	čəmən~čem	mən
2SG	tæ	štæ	štæ	štæ	tæ
3M	a	əja	jan	ja	ja
1PL	ama	čəba	čəba	?	ama
2PL	šoma/šəba	šəba	šəba	?	šoma
3PL	a	jon	əjon, jon	jon	jon
'who'	ki	?	čiki	čiki	ki, čiki

Table 5I: Chāli (maximum: 3-term, 1SG, 3SG.M/F, 3PL)

	Subject	Pres/Obj	Past/Obj	Poss	Past Agent
1SG	æz	čemen	čemen	čemen	men
2SG	tæ	eštæ	eštæ	eštæ	tæ
3M	æ	jæy	jæy	jæy	æy ~ ay ~ yi
3F	æyæ	jæyæ	jæyæ	jæyæ	æyæ ~ yæ
3PL	æy	jæyæ	jæyæ	jæyæ	æyo(n) ~ yo(n)
'who'	ki, ke	eške, eški		eške, eški	

5.3. The uses of the cases and Set₁/Set₂ PAMs in Tatic languages

Two-term case systems typically exhibit rather high polyfunctionality (Arkadiiev 2008). Tati dialects are a typical example of this phenomenon. Table 5J demonstrates the major functions of the Direct and Oblique cases. Various minor usages, e. g., Oblique marking of temporals (see ex. 109), have been omitted.

Table 5J: Uses of the cases and Set₁ Set₂ in Tatic

	Direct Case	Set ₁	Oblique Case	Set ₂
Subject, all Intransitives	+	+	-	-
Agent, Present system	+	+	-	-
Agent, Past system	(+) (§7.1.3.1) ¹	-	+(§7.1.3.1) ¹	+(§7.1.3.1) ²
Non-Salient Patient	+ (§7.1)	-?	-	-
Salient Patient, Present	-	-	+(§7.1.3)	+(§7.3) ¹
Salient Patient, Past	+ (§7.1.3) ¹	+ (§7.3) ³	± (§7.3)	± (§7.3)
Recipient/Indirect Object	-	-	+(§7.5)	+(§7.5.1) ⁴
Experiencer	-	-	+(§6.7)	+(§6.7)
Possessor/Genitive	-	-	+(§§)	+(§7.3)
Adjunct of Adposition	±	-	±	+
Goal (no Adposition)	+ (§9.1)	-	(+) ¹	-
Locative (no Adposition)	(+) ²	-	(+) ¹	-

1. occasional pattern; 2. for most lgs; 3. for some lgs; 4. minor pattern

5.4. Possessor constructions: reverse Ezafe, Ezafe, oblique case

Possessors generally precede the head in Caspian and most of Tatic using the REZ and the Oblique case, respectively. Ongoing Persian right-branching influence on Gilaki, however, has now created a split typology based on animacy: preposed possessors marked with the REZ are predominately animate, while many inanimate possessors are postposed and marked with the head-marking *Ezafe* construction.

5.4.1. Caspian: REZ and Ezafe marking of nominal possessors

Of particular interest are the results of a frequency count of the REZ vs. *Ezafe* with Animate vs. Inanimate nouns in three Caspian varieties: texts in Rashti (Rastorgueva 1971), Sāravi (Yoshie 1996), and Kordkheyli Mazanderani (Borjian and Borjian 2007a, 2007b, 2008). Rashti shows a preference for the REZ with animates (REZ: 90.1%) vs. the EZ with inanimates (REZ: 87.5%). Mazanderani, however, presents quite a different picture. While there is a clear preference for the REZ with both animate nouns (REZ: Sāravi: 97.7%/Kordkheyli 100%) and inanimates (REZ: Sāravi: 65.2%/Kordkheyli: 87.5%), the use of the EZ with inanimates seems to be only incipient here.

Rashti

- (05) Animate Possessors: Reverse *Ezafe* (Possessor-HEAD)

məryám-ə per iskəndár-ə harf-án=a gúš kun-è.
 P.N.-REZ father P.N.-REZ speech-PL=RA ear LV-3s
 ‘Maryam’s father listens to Iskandar’s words.’ (RR: 273, 118)
- (06) Inanimate Possessors: *Ezafe* (HEAD-possessor)

abul púl-ə nuyrə az xu refey akbər fa-gíft-ə.
 P.N. money-EZ silver from self friend P.N. PVB-take:PT-3s
 ‘Abul took silver coins from his friend Akbar.’ (RR: 264, 21)
- (07) Inanimate Possessors: Reverse *Ezafe* (Possessor-HEAD)

ná-tanəst-i-Ø xu dil-ə gáb=a iskəndár=a bá-g-ə.
 NEG-can:PT-IMF-3S self heart-REZ speech=RA P.N.=RA PU-say-3S
 ‘She couldn’t tell Iskandar her heart’s desires (lit: words)’ (RR: 277, 219)
- (08) Lāhijāni

Animate Adjunct (REZ)	Inanimate Adjunct (REZ)
<i>i asp-ə sum</i>	<i>tu istidad-ə zabon=æ da-n-i.</i>
this horse-REZ hoof	you talent-REZ language=RA have-PR-2S
‘this horse’s hooves’ (NJ,47)	‘You have language talent.’

(09) Sāravi

ša-e xæzun-e bar=e Ø-værd-e.
 shah-REZ treasure-REZ load=RA IMF-take:PT-3S
 ‘He would take the load of the Shah’s treasure (to the bazaar).’ (SY: 68)

dæ-ket-e gol-e báy-e dele.
 PVB-fell-3s flower-REZ garden-REZ in
 ‘She fell into a bed of roses (lit: a flower garden).’ (SY: 58)

5.4.2. Tatic: case marking of nominal possessors

Tatic varieties also have an adjunct-marking pattern in which an Oblique case-marked Possessor precedes the head. While Tatic may have some *Ezafe* tokens with postposed Possessors, there is no clear animacy-based split here (but see Vafsi below).

(10) Leriki

hakim-i söhbæt bæ-pe čok bi-bu-Ø.
 P.N.-os talk PR-must good SBJ-be-3S₁
 ‘Hakim’s talk should go well.’

sef-i donæ, zəməston-i donæ
 apple-os seed watermelon-os seed
 ‘apple seeds, watermelon seeds’

(11) Asālemi

čin-i ša ešte pi-ær nom če=yye?
 China-os shah you:POSS father-os name what-COP:3S₁
 ‘the king of China’ ‘What is your father’s name?’
 (ASA: 88)

(12) Khoini (Central Tati enclave)

pæs-é pæšmæ æ killeg-á nom či=æ?
 sheep-os wool that girl-OF name what-COP:3S₁
 ‘the sheep’s wool’ (YX: 169) ‘What is that girl’s name?’

The Harzani Oblique case is restricted to Possessor-marking (cf. Alignment, §7.1.1):

(13) hæsæn-i kætaeb

P.N.-os book
 ‘Hassan’s book.’ (AAK:108)

badam-i raz

almond-os orchard
 ‘almond orchard.’ (AAK:108)

5.4.3. The Vafsi possessive constructions

Vafsi has three main strategies with one subtype available for adnominal possession:

- 1) In the *Preposed Genitive Construction I*, used mostly for animate possessors, the adjunct noun/pronoun is in the Oblique case and precedes the head noun:

- (14) Animate nouns (predominant)

yey	tajere-y	bær	dokán-i	v-e.
one	merchant-OM	door	shop-OM	COP:PT-3S ₁
'It was a door of a merchant's shop.'				

- Inanimate nouns (minor)

bær-i	taxtæ	or-éysdæ=v-e.
door-OM	plank	PVB-arose=AUX:PT-3S ₁
'The door plank had come up.'		

Preposed Genitive Construction II: The above construction is occasionally reinforced by an enclitic possessive after the head noun, yielding a circumposed possessive:

- (15) zell-i bera=s b-áw-e.

woman-OF ₂	brother=3S ₂	PU-come:PT-3S ₁
'The woman's brother came over.'		

- in kelj-i dél=es bæ-šg-i-a-Ø.

this girl-OF ₂	heart=3S ₂	PU-break-PASS-PT-3S ₁
'This girl's heart broke.'		

- 2) The Right-branching "Vafsi *Ezafe*" and postposed nominal adjuncts

Reminiscent of Kurmanji, the Vafsi right-branching *Ezafe* Construction is both head- and adjunct-marked: the head has an *Ezafe* suffix followed by an Oblique-marked adjunct, but the *ezafe* itself is invariable and lacks the gender distinctions of Kurdish:

- (16) bædbæxti-e vóws-i róbb-e kudu-é

misfortune-EZ	Vafs-OM	paste-EZ	pumpkin-OF
'the misfortune of Vafs'			'pumpkin butter'

- 3) The Postposed Genitive Construction, generally found with non-human (mostly inanimate) adjuncts, has a postposed possessor in the Oblique case but lacks an *Ezafe*:

- (17) lášæ ešgár-i čera mašín-i bùæ kæbáb-i
 carrass deer-OM light car-OM smell kebab-OM
 ‘deer carcass’ ‘car headlights’ ‘the smell of kebab’

5.4.4. Other Tatic strategies

In Alviri, to the northeast of Vafsi, possession is exclusively right-branching with an invariable *Ezafe* suffixed to the head noun linking it to the postposed Possessor:

- (18) Alviri (YA: 179)
pur-e hæsæn e-væz-æ. *tete-y hæsæn e-væz-iæ.*
 son-EZ P.N. DUR-run-3SM₁ daughter-EZ P.N. DUR-run-3SF₁
 'Hassan's son is running.' 'Hassan's daughter is running.'

Juxtaposition: left-branching, Ø-marking of both Head and Adjunct

- (19) Razajerdi
dar šax *qæmbær ærōsi*
 tree branch P.N. wedding
 ‘tree branch, branch of the tree’ ‘Qambar’s wedding’

The Possessor precedes the head in Rudbāri and in some Tāleqāni varieties. The Rudbāri possessor is marked with the REZ while Tāleqāni has a marker that resembles the Oblique case form of other Tati varieties, but it is used agglutinatively in the plural and with pronouns, which indicates it is more likely a Reverse *Ezafe*:

- (20) Tāleqāni: Owrāzāni (AA1: 87)
boz-an-i *šir* *pa-i* *bon*
 goat-PL-REZ milk foot-REZ bottom
 'the goats' milk' 'sole of the foot'

Gurāni			
<i>mán-i</i>	<i>xanæ</i>	<i>gusfænd-i</i>	<i>sær</i>
I-REZ	house	sheep-REZ	head
‘my house’		‘the sheep’s head’	

- (21) Rudbāri: Tutkāboni
 i dár-e bul-an ayoz-e pust
 this tree-REZ branch-PL walnut-REZ skin
 ‘this tree’s branches’ ‘walnut shell’

Jubani	<i>gorg-e</i>	<i>zay</i>	<i>šiær-e</i>	<i>nom</i>
	wolf-REZ	child	husband-REZ	name
	‘wolf(’s) cub’		‘husband’s name’	

5.5. Adjectival modification

5.5.1. The reverse Ezafe, the Ezafe, and Ø marking (juxtaposition)

Jahāngiri (2003: 52) states that no linker morpheme is used in Lāhijāni with adjectives, but many of his examples are exceptions to this rule in that they use the REZ.

- (22) Lāhijāni, left-branching, no REZ

<i>pəsəroná</i>	<i>juraf-on</i>	<i>xoróm</i>	<i>zak</i>	<i>kujá</i>	<i>pəsár</i>
boy:ADJ	sock-PL	good	boy	small	son
‘boys’ socks’ (NJ: 109)		‘good child’ (NJ: 62)		‘younger son’ (NJ: 143)	

but with REZ:

- (23) *sərd-ə* *áw* *nər-ə* *gúrg* *kàl-ə* *mivá*
 cold-REZ water male-REZ wolf unripe-REZ fruit
 ‘cold water’ ‘male wolf’ ‘unripe fruit’
 (NJ: 162) (NJ: 104) (NJ: 145)

The right-branching *Ezafe* construction also occurs with adjectives in Lāhijāni. (In actuality most of Jahāngiri’s example sentences show Head-EZ Adjective). Sāravi is more consistently Adjective-HEAD, with a Reverse *Ezafe* on the adjunct:

- (24) Sāravi, left-branching (SH: 148)

<i>inje</i>	<i>čæn=ta</i>	<i>gæt-e</i>	<i>dikun</i>	<i>va-bæ-yy-e.</i>
here	a.few=UNC	big-REZ	shop	PVB-become-PT-3S ₁
‘A few big shops opened up here.’				

Lāhijāni, right-branching (NJ: 103)

<i>i</i>	<i>dirəxt-ə</i>	<i>jəngəl=e.</i>
this	tree-REZ	forest=COP:3S ₁
‘This is a forest tree’		

Vafsi

The adjective in Vafsi is simply juxtaposed after the head noun. In the Direct singular, the adjective agrees in gender with the noun, but plurality and other cases are only marked once, on the adjective, while the head noun remains invariable:

- (25) Masculine: ‘new broom’

Direct Sg.	Oblique Sg.	Direct Pl.	Oblique Pl.
<i>sizdæ</i>	<i>nu</i>	<i>sizdæ</i>	<i>nú-i</i>
broom	new	broom	new-OM

Direct Sg.	Oblique Sg.	Direct Pl.	Oblique Pl.
<i>sizdæ</i>	<i>nú</i>	<i>sizdæ</i>	<i>nú-e</i>
broom	new	broom	new-DP

Direct Sg.	Oblique Sg.	Direct Pl.	Oblique Pl.
<i>sizdæ</i>	<i>nu</i>	<i>sizdæ</i>	<i>nu-án</i>
broom	new	broom	new-OP

- (26) Feminine: ‘new shovel’

Direct Sg.	Oblique Sg.	Direct Pl.
<i>fí-æ</i>	<i>nú-æ</i>	<i>fí-æ</i>
shovel-DF	new-DF	shovel-DF
	<i>fí-æ</i>	<i>nú-e</i>
	shovel-DF	new-DP
	Oblique Pl.	
	<i>fí-æ</i>	<i>nu-án</i>
	shovel-DF	new-OP

Adjectival modification occasionally uses the Vafsi *Ezafe*, including Oblique marking of the adjunct. This pattern clearly must be a recent copy from Persian since Vafsi has not innovated a way for the *Ezafe* to be combined with feminine head nouns.

- (27)
- tæbæqe-ye zæif-i*
- in dús-d-e mehræbán-i=m=e*
- .
-
- class-EZ weak-OM this friend-EZ kind-OM-1S
- ₂
- =COP:3S
- ₁
-
- ‘the weak(er) class (of society)’ ‘He is my kind friend.’ (VF: B5,3)

5.5.2. General linker -æ: adjectival modification and noun-noun compounds

The unstressed General Linker -æ, suffixed to the preposed adjective or noun adjunct is a feature common to most of Tatic but is an areal feature that also encompasses Caspian languages. (See also §5.1.1, point 6 for the Caspian forms).

- (28) Leriki

Adjectival modificaton

<i>šàt-æ</i>	<i>dó</i>	<i>bìž-æ</i>	<i>nečí</i>
crooked-LINK	tree	sly-LINK	wolf
‘crooked tree’		‘crafty wolf’	

Possessive or Noun Compounding

<i>kàg-æ</i>	<i>pór</i>	<i>bìž-æ</i>	<i>mú</i>
hen-LINK	feather	goat-LINK	hair
‘chicken feathers’		‘goat hair’	

- (29) Asālemi (ASA: 90)

<i>sàr-æ</i>	<i>áesb</i>	<i>čák-æ</i>	<i>zué</i>	<i>màs-æ</i>	<i>patíl</i>
red-LINK	horse	good-LINK	son	copper-LINK	pot
‘red horse’		‘good son’		‘copper pot’	
<i>síf-æ</i>	<i>dár</i>				
apple-LINK	tree				
‘apple tree’					

(30) Māsulei

<i>rùk-æ</i>	<i>berá</i>	<i>pill-æ</i>	<i>berá</i>
small-LINK	brother	big-LINK	brother
‘younger brother’		‘older brother’	
<i>ræzin-æ</i>	<i>galəš</i>	<i>bərz-æ</i>	<i>kis-on</i>
rubber-LINK	galosh	rice-LINK	sack-PL
‘rubber galoshes’		‘rice sacks’	

(31) Kalāsuri

<i>koy-æ</i>	<i>usmún</i>	<i>miòn-æ</i>	<i>xó</i>	<i>jòwz-æ</i>	<i>dár</i>
blue-LINK	sky	middle-LINK	sister	walnut-LINK	tree
‘blue sky’ (KL: 272)		‘middle sister’		‘walnut tree’	

Razajerdi

<i>æsif-æ</i>	<i>deréxt</i>
apple-LINK	tree
‘apple tree’	

Vafsi also has this linker but it is no longer productive and is only found in frozen forms: *dèlæ-dárd* ‘stomach ache’, *néræ xér* ‘male donkey’.

5.6. Singular and plural vs. definiteness categories

Generic and Indefinite nouns are generally not pluralized but remain in the base form. Also see Numeral Classifiers (§5.8) for another use of the singular form of nouns. (See also sentences 9, 47, 86, 89, 129, 174, 183, 184, and 266 for more examples.)

(32) Sāravi

<i>haemíšeg</i>	<i>bakele=o</i>	<i>mærji</i>	<i>xær-im-be.</i>	<i>Lāhijāni</i>
always	bean=and	lentil	buy-PR-1s ₁	<i>sib</i>
‘I always buy beans and lentils.’				<i>bó-kašt-im.</i>
(SY: 139)				apple PU-plant:PT-1P ‘We planted apples.’

(33) Leriki

<i>kag</i>	<i>kižæ</i>	<i>bárd-æ-Ø.</i>
chicken	chick	carry-LOC-3s ₁
‘It carries off chickens and chicks.’		

Vafsi

<i>vəws</i>	<i>sí-æ</i>	<i>dir-e,</i>	<i>viz</i>	<i>dir-e.</i>
P.N.	apple-DF	have-3s ₁	walnut	have-3s ₁
‘Vafsi has apples, it has walnuts.’				

Generally only definite count nouns are pluralized in these languages:

- (34) Lāhijāni

Indefinite/Generic

pərviz xəyli gusfənd da-n-ə.
P.N. much sheep have-PR-3S
'Parviz has lots of sheep.' (NJ: 130)

Definite

gusfənd-on=a dū-kun təvile men.
sheep-PL=RA PVB-do stable:EZ inside
'Put the sheep in the stable.' (NJ: 153)

- (35) Sāravi

zəmin=re kel há-kon, bæ:d tim dák-šend.
land=RA plow PVB-do then seed PVB-sow
'Plow the land first then sow seeds.' (SH: 148)

gosfen-a=re vær-n-e čera.
sheep-PL=RA take-PR-3S pasture
'He takes the sheep to pasture.' (SH: 149)

- (36) Leriki

muzikant=æn hïst b-e?
musician=also EXIST COP:PT-3S₁
'Were there musicians, too?'

æ:la b-in čæy muzikant-on.
great COP:3P₁ he-poss musician-PL
'They were great, his musicians'

5.7. Adpositions

Caspian and most Tatic languages basically have left-branching NPs, including a predominance of postpositions. Given the heavy bilingualism with Persian and the socio-economic impact from Tehran throughout the Caspian, however, many right-branching patterns, such as prepositions, are entering these languages and are clearly becoming integrated into their grammars.

- (37) Postpositions in Lāhijāni

hasən bay-ə mien ísa-Ø. ali deraxt-ə sər ísa-Ø.
P.N. garden-EZ inside BE₃,3S P.N. tree-EZ on BE₃,3S
'Hassan is in the garden.' 'Ali is up in (lit: on) the tree.'

- (38) Prepositions commonly used in Lāhijāni
jujə lá-ye pələ hánna. xa bá-dar-i ta paiz.
 chick within-EZ pilaff BE₅ must PU-hold-2s until autumn
 ‘There is chicken in the pilaff.’ ‘You must keep (it) until autumn.’
 (NJ: 47)
- (39) Sāravi Postpositions (SY: 62)
e-m-be šeme nal-e sær.
 come-PR-1s you.PL:POSS porch-EZ on
 ‘I’ll come onto your veranda.’
- b-úr-em æt-ta kar-e dembal.
 PU-go-1s one-UNC work-EZ behind
 ‘Let me go after some job.’
- (40) Kafteji Postpositions
čél-æ ke dəlæ neh-áya.
 spindle-DF house inside BE₅.3SF₁
 ‘The spindle is in the house.’
- æm til kovrá=ræ b-om-é?
 this mud where=from PU-come:PT-3S₁
 ‘Where did this mud come from?’
- (41) Dikin Marāqeи Postpositions
séng=em-an dæ-fü-an cah-áe mion.
 stone-1S₂-FEM PVB-threw-3SF₁ well-OF inside
 ‘I threw a stone (f.) into the well.’
- æsif=em-an a-gáti-an zemin=de.
 apple-1S₂-FEM PVB-took-3SF₁ land=from
 ‘I got an apple (f.) off the ground.’

5.7.1. Mixed adpositional typology

Vafsi is located in an intersection zone between a postpositional isogloss to the north and a prepositional isogloss in the south (see Stilo 2009, 2012, see also Haig, this volume, chapter 2.3, §3, on similar phenomena across Northern Kurdish). It has a very mixed adpositional typology: postpositions, prepositions, circumpositions, as well as doublets (ex. 42), that is, one adposition that doubles as postposition or preposition, e. g., Locative vs. Allative, respectively:

(42) Locative (Postposition)

kúg-æ otaq=dæ æd-oss-e.
 cat-DF room=at DUR-sleep-3S₁
 'The cat sleeps in the room.'

Allative (Preposition)

kúg-æ báe-sse-Ø dæ benin otaq.
 cat-DF PU-go:PT-3S₁ to inner room
 'The cat went into the back room.'

(43) (Circumposition)

kúg-æ [sær] miz-i=dæ hott-e
 cat-DF on table-OM=at sleep:STAT-3S₁
 'The cat is asleep on the table.'

(Compound Preposition)

kúg-æ báe-væst-e [o sær] miz-i.
 cat-DF PU-go:PT-3S₁ to on table-OM
 'The cat jumped onto the table.'

The examples below show Leriki Prepositions, Postpositions, and Circumpositions:

- (44) *sabir væy b-e dī štæ žen-i.*
 P.N. there COP:PT-3S₁ with self wife-OM
 'Sabir was there with his wife.'

čoštæ=m pat=e bo ayl-on.
 breakfast=1S₂ cook:PT=AUX for child-PL
 'I cooked breakfast for 'the guys'.'

- (45) *tī máštæ mæktæb=aedæ=ž bæ-b-e?*
 you tomorrow school=at=2S₁ ALL-be-INF
 'Will you be at school tomorrow?'

kæ-y kum=o e-vašt-im.
 house-OB roof=from PVB-jump:PT-1S₁
 'I jumped off the roof.'

- (46) *æv [či] kæ-y=ku be-š-e-Ø.*
 he from house-OB=from PVB-go:PT-3S₁
 'He went out of the house.'

æv [či] šæhær=o omæ-Ø.
 he from city=from come:PT-3S₁
 'He came from the city.'

5.7.2. Adpositional polysemy

Note that in many Tatic and Caspian languages certain basic adpositions are polysemous, occasionally even with opposite meanings, e. g., Ablative-Allative:

- (47) Sāravi

=je Allative
hæme jur kar=je Ø-resi-ne.
 all kind work=JE IMF-arrive:PT-3P
 ‘They would get to all kinds of work/
 They would see to all ...’ (SY: 58)

=je Ablative
sere=je dær n-uri.
 house=JE out NEG-come:IMPER
 ‘Don’t come out of the house.’ (SY: 66)

- (48) =je Comitative

hæsen hosæyn=je košti ger-n-e.
 P.N. P.N.=JE wrestling take-PR-3S
 ‘Hassan wrestles with Hosseyn.’ (SH: 138)

- (49) Karani (YSFN)

=ku Locative
xav-e=m=ku do glæ=m æsp b-ind.
 sleep-OM=1S₂=KU two UNC=1S₂ horse PU-see:PT
 ‘I saw two horses in my dream.’

=ku Ablative
pe=m Māsāl-e=ku a-gerdes.
 father=1S₂ P.N.-OM= KU PVB-wander:PT
 ‘My father returned from Māsāl.’

- (50) =ku Allative

dé-ček-amen=eš kæ bar-e=ku.
 PVB-stick-CAUS=3S₂ house door-OM=KU
 ‘Stick it on the door of the house.’

Ø-marking of Allative
i ruj be-še-Ø ku.
 one day PU-go:PT-3S₁ mountain
 ‘Once he went to the mountain.’

On the other hand, the same sense may be expressed by various synonymous adpositions, often with the same adpositions that are polysemous.

(51) Leriki

iṁ kæ kærpič=ædæ=y.
 this house brick-^{ÆDÆ=COP:3S₂}
 ‘This house is (made) of bricks.’

iṁ kæ či kærpič=o=y.
 this house from brick=from=COP:3S₂
 ‘This house is (made) of bricks.’

5.8. Numeral classifiers

West Iranian languages have simple Numeral Classifiers systems, with only two or three common classifiers after numerals and the quantifier *čænd*, ‘a few; how many?’ (see also Stilo 2018a). These classifiers also involve a separate issue: the use of singular forms of nouns after numerals. Classifiers in this area consist of two basic types:

- 1) Universal Numeral Classifier (UNC) used for most nouns (including human):
 - a) Most Tati languages have the form *dane*, *danæ* ‘seed, grain/kernel’.
 - b) All Talyshi dialects and some neighboring Tati dialects use *g(ə)læ*, a UNC, also meaning ‘seed, grain’.
 - c) Caspian languages use only *=ta* ~ *=tæ* throughout, irrespective of the animacy of the noun or the head numeral/quantifier.
- 2) S. Talyshi and much of Tati also have the Human Numeral Classifier (HNC) *næfær* ‘person’, but the preferred choice for human nouns is still the UNC listed above. The HNC is rare in Caspian languages.

Numeral Classifiers are generally optional in most of the languages under discussion. In Caspian languages and probably most of Talyshi, however, the use of the UNC seems to be obligatory, even with human nouns. The numeral ‘one’ also functions as the indefinite article and the UNC is also commonly used with ‘one’ in this function.

(52) Lāhijāni: UNC *ta*

<i>do=ta</i>	<i>čere-ye</i>	<i>jaedid</i>	<i>iyn-əm.</i>	<i>yek=ta</i>	<i>lako</i>	<i>bá-ma-Ø.</i>
two=UNC	face-EZ	new	see-1s	one-UNC	girl	PU-come:PT-3s
‘I see two new faces.’				‘A girl came (here).’ (NJ: 52)		

(53) Sāravi: UNC *ta*

<i>in</i>	<i>se-ta</i>	<i>se*</i> .	<i>æt-ta</i>	<i>mærdi</i>	<i>bi-ye</i>	<i>bez</i>	<i>dašt-e.</i>
this	3-UNC	apple	one-UNC	man	was-3s	goat	had-3s
‘These three apples’				‘There was a man (who) had a goat.’			
(SH: 72) *(See §3.2 for phonology of vowels)				(SH: 72)			

In Leriki the UNC with the indefinite marker *-y* (derived from the numeral *i* ‘one’) may function without the numeral as an article:

- (56) *ištæ jif=o gilæ-y dæftær be-várd-æ-Ø.*
 self pocket=from UNC-INDF notebook PVB-bring:PT-LOC-3S_i
 'He takes a notebook out of his pocket.'

5.9. Reflexives

In Gilaki and nearby Tati, reflexive pronouns are only used in the 3rd person. SAP persons use the corresponding case forms of these pronouns (cf. Romance, German).

- (57) Lāhijāni
mu m-a bə i kar adát bə-da-m.
 I I-RA to this work habit PU-give:PT-1s
 'I've accustomed myself to this.' (NJ: 98)

(58) *tu t-a niga nú-kun!* *xíu=š=a bú-kušt-ə.*
 you you-RA look NEG-do self-3S₂=RA PU-kill:PT-3s
 'Don't look at yourself!' (NJ: 118) 'He killed himself.' (NJ: 117)

Mazanderani (*še*), C. Caspian (*xoštə*), and most Tatic (*VštVn*) exhibit one reflexive pronoun for all persons. Tatic varieties that have Set2 possessive enclitics, e. g., Vafsi, optionally attach them to the reflexive pronoun (cf. Eng. *my-self*, Pers. *xod-aem*).

- | | | |
|-------------|------------------------------|---------------------------------|
| (59) Sāravi | <i>še=re bæzek há-kan!</i> | Rāmsari (Central Caspian) |
| | self=RA adorn PVB-do | why self-RA hit-2s |
| | ‘Make yourself up!’ (SY: 56) | ‘Why are you hitting yourself?’ |

(60) Leriki

ištæn-í či tók=o ayirmiš bř-kæ-m,
 self-OB ABL current=ABL separated SBJ-do-1S₁
 ‘...in order to separate myself from the (electric) current, ...’

Keringāni and Harzani

K: *mun hišton yæ=me.*

H: *mæn ištæn yær=mæ.*
 I self hit:PT=1S₂

‘I hit myself.’ (AAK: 85, 107, resp.)

(61) Vafsi

(with Set ₂)	(without Set ₂)
<i>n-ím-aesdæ ešdén-i=s báe-zæn-e. æz ešdén-i r-vin-om.</i>	
NEG-1S ₂ -let:PT self-OM=3S ₂ PU-hit-3S ₁	I self-OM DUR-see-1S ₁
‘I didn’t let him hit himself.’	‘I see myself.’

Māsulei has mixed typology using both of the above patterns for different functions:

- Direct Objects, all persons: dedicated Reflexive (no Set₂ enclitics)
- Adpositions, Possessive, 3rd SG/PL: dedicated Reflexive (no Set₂ enclitics)
- Adpositions, Possessive, 1st/2nd SG/PL: Genitive case of SAP Pronouns

(62) *æz be øštæn=ni bá-kæš-əm. tø øštæn mæ-kæš.*
 I must self=also SBJ-kill-1S₁ you self NEG:IMPER-kill
 ‘I must kill myself, too.’ Don’t you kill yourself.’
 (LM: 38; I: 16) (LM: 38; I: 16)

(63) *øčəra čəmən kellæ dæ-benn-əm? a-n øštæn kæ.*
 why I:POSS daughter PVB-tie-1S₁ come-3P₁ self house
 ‘Why should I tie my daughter up?’ ‘They come to their house.’
 (LM: 46, II, 33) (LM: 52, IV, 1)

6. Verbs: initial generalities

This section introduces Present vs. Past verb stems, lexical preverbs, negation, and the formation, diachrony, and usages of the various tenses.

6.1. Present vs. past verbal systems

The set of paradigms based on the present stem (Present, Subjunctive, Imperative) are called the Present system below and those based on the past stem (Preterit, Perfect tenses, and usually Imperfect) are called the Past system. The number and

members in each vary by language, e.g., the Talyshi Imperfect belongs to the Present system.

This division into two systems of tenses is purely morphological and is irrelevant for Caspian and Tatoid, but for Tatic it establishes a tense-based split in Intransitive and Transitive conjugations. This split affects the choice of PAMs (§7.3) and ultimately the alignment of core arguments (§7.1.3.1): Leriki *æv vít-ə* (**he:DIR** run:PT-3S₁) ‘s/he ran’ vs. *æy vít-ič=e* (**he:OBL** sift:PT=3S₂=AUX) ‘s/he sifted’. The Present alignment in Tatic is Nominative-Accusative and the Past system may have Ergative alignment or some vestige of original Ergativity.

6.2. Present and past stem formation

West Iranian languages use multiple strategies for forming the past stems of verbs and no fixed rules can be set up to derive the past stem from the present. Some languages have a high preponderance of *-d/-t* suffixes in the past stem, often accompanied by other, perhaps substantial, changes in the present root. Table 6A presents a small sampling of present and past stems in various Caspian and Tatic languages:

Table 6A: Present and past verb stems in Caspian and Tatic

Lexeme	Rästi	Tonkb	Särvä	Koluri	Harzan	Kering	Kajal	Asälm	Mässäl	Kafteji	Chäli	Alvir	Vafsi
bring	<i>avər</i>	<i>hær</i>	(y)ar	<i>var</i>	<i>æ(r)</i>	<i>er</i>		<i>vær</i>	<i>ar</i>	<i>or</i>	<i>ar</i>	<i>or</i>	<i>ar</i>
past	<i>avərd</i>	(y)ärd	ard	<i>vard</i>	<i>ord</i>	<i>vard</i>	<i>avärd</i>	<i>värd</i>	<i>ard</i>	<i>ord</i>	<i>ard</i>	<i>ord</i>	<i>ard</i>
fall	<i>kəf</i>	<i>kæ</i>	<i>kaef</i>	<i>gen</i>	<i>gen</i>	<i>gin</i>		<i>gən</i>	<i>läk</i>	<i>gen</i>	<i>gen</i>	<i>gen</i>	<i>gen</i>
past	<i>käft</i>	<i>kæt</i>	<i>kæt/ket</i>	<i>geness</i>	<i>genesd</i>	<i>ginest</i>	<i>kkæt</i>	<i>gənəst</i>	<i>läki</i>	<i>kæt</i>	<i>kæt</i>	<i>kæt</i>	<i>kætt</i>
sleep	<i>xus</i>	<i>xas</i>	<i>xas</i>	<i>xesa</i>	<i>hes</i>	<i>fes, het</i>	<i>hös</i>	<i>xəs</i>	<i>xəs</i>	<i>xos</i>	<i>u-xos</i>	<i>xös</i>	<i>hoss</i>
past	<i>xuft</i>	<i>xat</i>	<i>xat</i>	<i>xet</i>	<i>het</i>	<i>fesi,het</i>	<i>höt</i>	<i>xət</i>	<i>xət</i>	<i>xott</i>	<i>u-xott</i>	<i>xöt</i>	<i>hott</i>
hit	<i>zən</i>	<i>zæn</i>	<i>zæn</i>	<i>zæn</i>	<i>yæn</i>	<i>yæn</i>	<i>žæn</i>	<i>žæn</i>	<i>žæn</i>	<i>zæn</i>	<i>zæn</i>	<i>zæn</i>	<i>zen</i>
past	<i>ze</i>	<i>zæ</i>	<i>zu</i>	<i>zi</i>	<i>yær</i>	<i>yi, žiæ</i>	<i>že</i>	<i>že</i>	<i>žænd</i>	<i>zzə</i>	<i>zæ(n)d</i>	<i>zind</i>	<i>zæ</i>
say	<i>g</i>	<i>gu</i>	<i>ga</i>	<i>vaj</i>	<i>ös</i>	<i>vut,vož</i>	<i>vaj</i>	<i>vaj</i>	<i>vat</i>	<i>vaj</i>	<i>(v)aj</i>	<i>vaj</i>	<i>vaz</i>
past	<i>guft</i>	<i>gut</i>	<i>gut/gat</i>	<i>vat</i>	<i>öt</i>	<i>vut/vot</i>	<i>h/vat</i>	<i>vat</i>	<i>vat</i>	<i>vat</i>	<i>(v)at</i>	<i>vat</i>	<i>vatt</i>
take	<i>gir</i>	<i>gir</i>	(g)ir	<i>gir</i>	<i>ge-n</i>	<i>gi(r)</i>	<i>gi(r)</i>	<i>ger</i>	<i>ger</i>	<i>gir</i>	<i>u-ger</i>	<i>gir</i>	<i>gir</i>
past	<i>gift</i>	<i>git</i>	(g)it	<i>gæt</i>	<i>got</i>	<i>get(i)</i>	<i>gæt</i>	<i>gæt</i>	<i>get</i>	<i>gæ(rt)</i>	<i>geræt</i>	<i>giræt</i>	<i>girætt</i>

In N. Talyshi and some N. Tati there is a merger of many Present and Past stems. Progressing away from the “epicentre” of this isogloss (N. Talyshi-Kaläsuri/Khoynarudi) westward through Keringani to Harzani and southward along the Talyshi dialect chain, one gradually finds more distinct present vs. past verb roots until we reach C. and S. Talyshi which fully retain these distinctions. Examples:

Table 6B: One-stem and two-stem verbs in N. Talyshi, N. Tati

	1 stem	2 stems
‘cook’	Lr, Ab: <i>pat</i> Ks: <i>pat</i> ; Xn: <i>pæt</i>	Jk: <i>pej~ž/pæt</i> ; Kr: <i>paj/pat</i> ; Hz: <i>pæy/pot</i>
‘run’	Lr, Vz: <i>vit</i> ; Sy: <i>vit</i> ; Xn: <i>vrit</i>	Jk: <i>(y)vij/vit</i> ; Kr: <i>ver/veri</i> ; Hz: <i>veriž/veret</i>
‘say’	Lr, Ab, Vz: <i>vot</i> Xn: <i>vut</i> ; Ks: <i>ut</i>	Jk: <i>(v)aj/(v)at</i> ; Kr: <i>vož/vot</i> ; Hz: <i>ös/öt</i>
‘sell’	Lr: <i>havat</i> ; Vz: <i>hæt</i> ; Ks: <i>herat</i>	Jk: <i>hæš/hæt</i> ; Kr: <i>feraš/ferat</i>
‘sit’	Lr: <i>ništ</i> ; Ab: <i>nešt</i> ; Xn: <i>nešt</i>	Jk: <i>nəš/nəšt</i> ; Ks: <i>neš/nešt</i> ; Kr/Hz: <i>nešin/nešd</i>

(Lr: Leriki, Ab: Anbarani, Sy: Sayyadlar, Vz: Viznei, Jk: Jowkandāni, Ks: Kalāsuri, Xn: Khoynarudi, Kr: Keringāni, Hz: Harzani)

Even the dialects at the center of this isogloss keep the distinction in a few verbs: (Leriki/Anbarani/Viznei) *žæn/žæ* ‘hit’; (Leriki/Sayyādlar/Viznei) *hæ/hard* ‘eat’; *stæn/sæ* ‘buy, take’; (Jowkandāni) *stun/stæ* ‘buy, take’; (Kalāsuri) *estæ/estær* ‘buy, take’, *hær/hard* ‘eat’; (Keringāni) *ta/ti* ‘buy, take’, *fa(r)/hard* ‘eat’, *yæn/yi* ‘hit’; (Harzani) *asda(n)/astar* ‘buy, take’, *hæn/hord* ‘eat’, *yæn/yær* ‘hit’. (See also §4.2.1, point 14)

6.3. Lexical preverbs

While preverbs in Tatic and Caspian languages can have a purely directional sense, they more often serve to create unpredictable lexical extensions or total semantic shifts in a verb. Lexical Preverbs in Tatic, depending on the language and the tense, may be separated from the root by mesoclitic Set₂ PAMs encoding the A (and in some cases the P). Languages with prefixed durative TAMs (*me-*, etc.) or Allative TAMs that allow preverbs in durative tenses generally place the preverb before the TAM:

(64) Leriki

<i>o-b-æ-gard-é=m</i>	<i>o-æ-gard-í-m</i>
PVB-FUT-DUR-go.around-INF=1S ₁	PVB-DUR-go.around-IMF=1S ₁
‘I will return’ (Future)	‘I would return’ (Imperfect)

(65) Kafteji

<i>a-m-xšuj-əm.</i>	<i>a-y-vendær-emæ.</i>
PVB-DUR-knead-1S ₁	PVB-DUR-stand-1S ₁
‘I knead.’	‘I stand, stay.’ (YA: 183)

In the Present in Asālemi and Keringāni the preverb attracts the Set₁ PAMs (plus

the AUX in Keringani) leftward and it also suppresses the Present *bə-* (cf. Leriki Future):

(66) Asālemi

<i>b-æ-gærdest=im ></i>	<i>o=m-æ-gærdest.</i>
PR-DUR-wander:PT=1S ₁	PVB=1S ₁ -DUR-wander:PT
'I go around'	'I return'

Keringāni

<i>be-kard=eš=e ></i>	<i>u=š=e-kard.</i>
PR do:PT=2S ₁ =AUX	PVB=2S ₁ =AUX do:PT
'you do'	'you open'

Preverbs with motion verbs usually yield quite transparent lexical results. In the case of verbs that have no intrinsic directional sense, however, either the preverb, the verb root, or both may become somewhat semantically bleached:

Table 6C: Lāhijāni lexical preverbs

Base form	Meaning	Derived forms with Preverbs	
<i>xut-</i>	sleep	<i>fu-xut-</i>	hide (intr.)
		<i>du-xut-</i>	lie in ambush
		<i>də-xut-</i>	shrink (of a balloon)
		<i>va-xut-</i>	wake up suddenly, start
<i>git-</i>	catch, grab	<i>vi-git</i> [> vi-t]	pick up
		<i>va-git</i> [> væ-yt]	light, kindle
		<i>ha-git</i> [> hæ-yt]	get, buy
<i>če-</i>	pick, pluck	<i>fi-če</i>	gather up
		<i>də-če</i>	set up, pile up
<i>kašt-</i>	plant, sow	<i>fə-kašt-</i>	line up (trans.)

Vafsi (three preverbs: *dær-*, *or-*, *har-*); *keš-/keša-*

- (67) *b-ím-keša* (No preverb) *ór=om-keša*
 PU-1S₂-pull:PT PVB=1S₂-pull:PT
 'I pulled, dragged' 'I weighed; I drew (water from well)'

- (68) *hár=om-keša* *dáér=om-keša*
 PVB=1S₂-pull:PT PVB=1S₂-pull:PT
 'I ground (grain, etc); 'I struck (match); I drew (line); I set up (wall);
 I flattened out (metal)' I tied up (bundle); I pulled (from middle of
 pile)'

Leriki (6 preverbs: dæ-, pe-, e-, o-, be-, sæ-)

Base Root	Infinitive	Meaning
<i>ništ</i>	<i>ništ-e</i>	'to sit'
+ <i>pe-</i>	<i>pe-ništ-e</i>	'mount (horse), get in/on (car, train, bicycle), ride'
+ <i>dæ-</i>	<i>dæ-ništ-e</i>	'to bathe (intrans.), swim (go bathing)'
+ <i>sæ-</i>	<i>sæ-ništ-e</i>	'to pursue, run down, drive out, chase, go after'

Gurāni (Tāleqāni): Base form: *keš* ~ *kəš-* (1st singular)

- (69) *mí-kəš-om* *há-mi-keš-om* *vé-mi-kəš-om*
 DUR-pull-1S₁ PVB-DUR-pull-1S₁ PVB-DUR-pull-1S₁
 'I pull' 'I stretch out' 'I pull on (shoes, etc.)'

 vá-mi-keš-om *dé-mi-keš-om*
 PVB-DUR-pull-1S₁ PVB-DUR-pull-1S₁
 'I spread/unfold' 'I serve up, dish out'

Many simplex verb roots occur with no preverbs and others never occur without one:

Vafsi: *nešin/nešd* (This verb root does not occur without a preverb).

- (70) *há-nešesde-ym* *ó-nešesde-ym* *dé-nešesdæ-Ø*
 PVB-sit:PT-1S₁ PVB-sit:PT-1S₁ PVB-sít:PT-3S₁
 'I sat down' 'I mounted (a horse)' 'it alit (of birds); settled'

6.3.1. Lexical preverbs with aspectual functions

Vafsi and Mazanderani only use lexical preverbs in the affirmative of punctual tenses.

Vafsi: *gir-/girætt-*

- (71) *b-ís-girætt.* (Preterit affirmative) *ór=es-girætt.*
 PU=3S₂-take:PT PVB=3S₂-take:PT
 'S/he caught, grabbed.' 'S/he picked up, lifted.'
- (72) *hár=es-girætt.* *dáér=es-girætt*
 PVB=3S₂-take:PT PVB=3S₂-take:PT
 'S/he bought, got (from someone), 'It started (to rain), (the fire) took;
 peeled (fruit, bark), closed (door)' (the current) got blocked up'

In the Present and the Imperfect and in all negative forms, the preverbs of these verbs are deleted and the one resulting verb form will have all the above meanings:

- (73) *aer-gir-e.* *n-ís-girætt.*
 DUR-take-3S₁ NEG=3S₂-take:PT
 ‘S/he catches, grabs, picks up, buys, ‘S/he didn’t catch, grab, pick up,
 gets, peels, closes, (fire) takes, buy, get, peel, close (door), etc.’
 etc.

6.4. Verbal negation

Negation in the area is usually a simple prefixing of a negator, *nV-* (with some vowel harmony but dissimilation in Mazanderani), which always attracts stress and also suppresses the punctual TAM *bV-*. Tati has a special Imperative negator, *mé* (§6.5.2).

- (74) Lāhijāni
xó-n-əm > nó-xo-n-əm *bó-he-m > ná-he-m*
 read-PR-1s NEG-read-PR-1s PU-buy:PT-1s NEG-buy:PT-1s
 ‘I read > I don’t read’ ‘I bought > I didn’t buy’

- (75) Sāravi
xor-m-e > náe-xor-m-e *Ø-xass-e > ná-Ø-xass-e*
 eat-PR-3s NEG-eat-PR-3s IMF-want-3s NEG-IMF-want-3s
 ‘s/he eats > s/he doesn’t eat’ ‘s/he wanted > s/he didn’t want’

- (76) Vafsi
aer-vin-óm > n-áer-vin-om *b-ím-díæ > n-ím-díæ*
 DUR-see-1S₁ NEG-DUR-see-1S₁ PU-1S₂-see:PT NEG=1S₂-see:PT
 ‘I see > I don’t see’ ‘I saw > I didn’t see’

The negator usually follows a lexical preverb, except for Māsulei where it precedes. In Vafsi and Mazanderani the preverb is deleted in negation (§6.3.1):

Lāhijāni	Koluri	Kelāsi	Chāli	Māsulei
<i>vi-n-gi-n-əm</i>	<i>pe-mé-gi</i>	<i>a-né-m-</i> <i>pær̩s=əm</i>	<i>a-mé-di</i>	<i>ná-pi-ger-əm</i>
PVB-NEG-take- PR-1S ₁	PVB-NEG-take	PVB-NEG-PR- ask-1S ₁	PVB-NEG-give	NEG-PVB-rise-1S ₁
‘I don’t pick up’	‘don’t pick up’	‘I don’t ask’	‘don’t give’	‘I don’t pick up’

Special cases in negation

In the Asālemi Present negative, the Allative *bə-* drops leaving the durative *-æ* and the negator attracts the AUX forward. The parallel N. Talyshi Future has two forms: 1) the negator is simply prefixed; 2) *æ-* + (Pres.) stem + Negative:AUX (Miller 1953: 157).

Asālemi Present

	Affirmative	Negative	
No Preverb	<i>b-æ-znóst=ím</i>	<i>nó=m-æ-znóst</i>	'I know'
Preverb	<i>dé=m-æ-pærəsəst</i>	<i>dæ-nó=m-æ-pærəsəst</i>	'I ask'

Lerik Future

Affirm.	<i>b-æ-dó=m</i>	'I will give'	<i>bæ-sé=m</i>	'I will get, buy'
Neg. I	<i>ní-bæ-do=m</i>	'I won't give'	<i>ní-bæ-se=m</i>	'I won't get, buy'
Neg. II	<i>æ-dé=ni-m</i>	'I won't give'	<i>æ-stáe(n)=ni-m</i>	'I won't get, buy'

In Talyshi and neighboring Tati, the analytic tenses usually place the negative AUX *after* the verb stem, an unusual situation within Iranian (see also Lerik Neg. II above):

(77) Leriki

<i>ički=aen</i>	<i>pe-šom-dæ</i>	<i>ni-š?</i>
spirits=also	PVB-drink-LOC	NEG:AUX-2S ₁
'And you don't drink?'		
<i>geymæt=ím</i>	<i>sæ</i>	<i>ni-Ø.</i>
grade-1S ₂	take:PPL	AUX:NEG-3S ₁
'I haven't gotten (my) grades.'		

(78) Māsulei

<i>æ</i>	<i>mærdek-é</i>	<i>vat-é</i>	<i>ni-æ.</i>
that	man-OB	say-PPL	AUX:NEG-3S ₁
'That man hasn't said (it).'			

Koluri

<i>híči=m</i>	<i>hærd-é</i>	<i>ni-Ø.</i>
nothing-1S ₂	eat-PPL	AUX:NEG-3S ₁
'I haven't eaten anything.'		

6.5. Verbs: tense formation, diachrony, areal features, and usages

6.5.1. Present tense

The various Present markers in Caspian and Tatic are: A) Ø marking; B) *mV*, *em-*; C) *æt*; D) Locative markers; E) Allative markers, and F) the *-Vn-* suffix. Types B – E are Durative markers used in any given language for both the Present and the Imperfect. See Progressives (§6.5.8.2) for another Present tense type in Charozh.

The Ø Present marker forms a small isogloss with languages of four groups: Western Gilaki, Southern Talyshi, Central Tati, and Rudbāri (Tatoid). It is one of

the main features distinguishing Western Gilaki from Eastern Gilaki. Tonekāboni and Khoini also exhibit this typology (and both have two Present tenses) but lie slightly outside the area of this isogloss.:

Contiguous languages belonging to this isogloss

	Rashti (Gilaki)	Māsulei (S. Talyshi)	Koluri (C. Tati)	Tutkāboni (Rudbāri)	Khoini (C. Tati)	Tonekāboni (C. Casp.)
1SG	Ø-dan-ám	Ø-zon-ám	Ø-zan-ám	Ø-dan-óm	Ø-tæj-em	Ø-xár-om
2SG	Ø-dan-i	Ø-zon-é	Ø-zan-i	Ø-dan-i	Ø-tæj-i	Ø-xár-i
3SG	Ø-dan-á	Ø-zon-á	Ø-zan-é	Ø-dan-é	Ø-tæj-ek	Ø-xár-e
	'know'	'know'	'know'	'know'	'run'	'eat'

6.5.1.1. Languages with the *mV-* present durative marker

The *mV-* prefix may be the same as Persian *mi-*, derived from a particle *hamē-* (< **hama-aiva-*, Windfuhr 1989: 256), but is unstressed in the languages of this area.

	Kajali	Hezārrudi	Kafteji	Dikin-Marāqeи	Eshtehārdi
1SG	me-mær-ém	mi-gen-ém	me-gæn-ám	me-zon-én	mi-væz-ém(æ)
2SG	me-mær-i	mi-gen-i	me-gæn-i	me-zon-i	mi-væz-iš(æ)
3M	me-mær-é	mi-gen-é	me-gæn-á	me-zon-á	mi-væz-á
3F	me-mær-íæ	mi-gen-íæ	me-gæn-éæ	me-zon-égan	--
	'die'	'fall'	'fall'	'know'	'run'

6.5.1.2. Languages with a prefixed *æt-* marker

Vafsi, Alviri, and Vidari have the durative marker *æt-*, forming an isogloss extending southwards to the Persian Gulf (Stilo 2007a: 108). It varies in Vafsi according to the following consonant. Alviri and Vidari drop the dental before a consonant.

Vafsi:	vowel	velar/labial	alveolar/ palatal	Alviri	Vidari
1SG	æt-ar-óm(e)	ær-gen-óm(e)	æd-dó-m(e)	e-væz-émæ	étt-a-m(e)
2SG	æt-ar-i	ær-mir-i	æn-nevis-i	e-va-y	é-ven-i
3SG	æt-ar-é	ær-vin-é	æ(č)-čin-é	et-ašin-á	é-ven-e
3F	--	--	--	et-ašin-íæ	--
	'bring'	'fall, die, see'	'give, write, pick'	'run, say, throw'	'come, see'

(See also §6.5.2.1 for the palatalization of roots in *s-/z-*: *saz-*, *zan-* > *æčazom*, *æjanom*)

6.5.1.3. TAMs of adpositional (locative/allative) origin

Locatives: Locative morphemes plus the infinitive of the verb, followed by a copula encoding tense and person in an analytic structure ('am/was in doing'), form general durative paradigms in N. Talyshi. In the Caspian group, Harzani, and Kalāsuri these forms have a Progressive sense (§6.5.8), separate from the Present/Imperfect tenses.

Northern Talyshi: Locative Present Tense (full and short forms)

	Full form -V	Full form -t/d	Short Form -t/d
	Stem + Stressed infinitive formant +		Stressed stem + Loc +
	Locative + Copula		Copula
1SG	š-é-dæ=m	vot-é-dæ=m	vót-dæ=m
2SG	om-é-dæ=ž 'go, come'	mand-é-dæ=ž 'say, stay'	mánd-dæ=ž > mónd-æ=ž 'say, stay'

Allatives: The Allative preposition *bV*, which is different from the Subjunctive marker, is an unstressed formant in another analytic Present/Past durative in Asālemi and all varieties of Northern Tati, as well as the Future in Northern Talyshi (§6.5.4).

	Harzani	Keringāni	Kalāsuri	Asālemi
1SG	<i>bə-het=en</i>	<i>bö-öt=en</i>	<i>be-het=en=e</i>	<i>be-het=em</i>
2SG	<i>bə-het=e</i>	<i>bö-öt=e</i>	<i>be-het=eš=e</i>	<i>be-het=eš</i>
3SG	<i>bə-het=ən</i> 'sleep'	<i>bö-öt=ən</i> 'say'	<i>be-het=i=e</i> 'sleep'	<i>be-het=i=ə</i> 'sleep'

6.5.1.4. Diachronic present participial formation

The Caspian Present formant (*V*)*n* (lacking in W. Gilaki), is traditionally derived from an old present participle through phonetic reduction: **ant* > *Vnd* > (*V*)*n(n)*. In Lāhijāni *ən* occurs after roots ending in most consonants, but *n* after a vowel. After a root-final *n-*, the Present marker is lost; after *r*, the latter is lost and *-n-* retained (but the *-r-* of the root reappears in the Subjunctive):

	Lāhijāni	Kelārdashti	Tonekāboni
	Consonant - <i>V</i>	- <i>n</i>	- <i>r</i> (xor-)
1SG	<i>xós-ən-əm</i>	<i>á-n-əm</i>	<i>dón-əm</i>
2SG	<i>xós-ən-i</i>	<i>á-n-i</i>	<i>dón-i</i>
3SG	<i>xós-ən-ə</i> 'sleep'	<i>á-n-ə</i> 'come'	<i>xó-n-ə</i> 'know'
			<i>bær-eme</i>
			<i>xár-en-om</i>
			<i>bær-n-i</i>
			<i>xár-en-i</i>
			<i>bær-n-i</i>
			<i>xár-en-e</i>
			'eat'
			'carry off'

Amini (2003) gives two Present tenses for Tonekāboni, one with the *-en-* suffix and one with the Ø marking presented above. It seems, however, that these two Present tenses are quite blurred in usage and often interchangeable:

- (79) *in či ke Ø-kár-im Ø-xár-im* – *ye číz-e*
 this what SUB plant-PR₁.1P eat-PR₁.1P one thing-EZ
ziádi=aem ruš-en-im.
 much=also sell-PR₂.1P
 ‘That which we plant, we eat—a lot of it we also sell.’ (Amini: 112)

The Mazanderani variant of *-Vn-*, simply *-n-* after roots in a single *r* or *l*, is lost in consonant clusters (**xornme*), and is *-en-* after roots in other consonants. When the nasal of the Present formant is followed by another nasal, there may be assimilation of the first nasal and dissimilation of the second: **xa-n-me > xa-m-be* 'I want'.

	Sāravi	Bābolsari
1SG	<i>kæfembe</i> < * <i>kæfenme</i>	<i>xorme</i> < * <i>xornme</i>
2SG	<i>kæfeni</i> < * <i>kæfeni</i>	<i>xorni</i> < * <i>xorni</i>
3SG	<i>kæfene</i> < * <i>kæfene</i> 'eat'	<i>xorne</i> < * <i>xorne</i> 'have'

Bābolsari (and others) generalizes resulting geminated nasals, Sāravi changes them to nasal + homorganic stop, while Āmoli (etc.) reduces the gemination:

	underlying	Bābolsari	Āmoli	Sāravi	Sāravi
	Present stems ending in Vowel				
1SG	* <i>dənme</i>	<i>demme</i>	<i>demə</i>	<i>dembe</i>	<i>dumbe</i> < * <i>dun-n-me</i>
2SG	* <i>dəni</i>	<i>denni</i>	<i>deni</i>	<i>deni</i>	<i>dundi</i> < * <i>dun-n-i</i>
3SG	* <i>dəne</i>	<i>denne</i>	<i>denə</i>	<i>dene</i>	<i>dunde</i> < * <i>dun-n-e</i>
	‘give’				
	‘know’				

Examples of Present tenses:

- | | | |
|------|---|---|
| (80) | Lāhijāni
čan saet xós-ən-i?
how.many hours sleep-PR-1s
'How many hours do you sleep?' | Sāravi
ve unta=re xa-n-e.
s/he that.one=RA want-PR-3s
'S/he wants that one.' (SH: 129) |
| (81) | Leriki
gofe lok-n-é-dæ-Ø.
cradle rock-CAUS-INF-LOC-AUX:3s ₁
'S/he is rocking the cradle.' | Vafsi
tæ če r-kær-i?
you what DUR-do-2s ₁
'What are you doing?' |

6.5.2. Subjunctive and imperative:

The Subjunctive/Imperative marker in West Iranian languages is usually described as a stressed prefix *bV-*. In examining this form more closely, however, I find that the *bV-* and the associated stress must be seen as separate phenomena for various reasons:

- a) Verbs with lexical preverbs suppress *bæ-* but retain the stress in initial position. In languages that have a Ø Present marker, therefore, the stress may be the only way to distinguish the Subjunctive from the Present when preverbs are present:

	Present	Subjunctive	
Rashti:	<i>fadám</i> (<i>fa-d-ám</i>)	<i>fádəm</i> (<i>fa-'d-əm</i>)	'I give'
Māsulei:	<i>adæm</i> (<i>a-dæ-ám</i>)	<i>ádæm</i> (<i>a-'dæ-əm</i>)	'I give'
Koluri:	<i>díæsám</i> (<i>di-æs-ám</i>)	<i>díæsəm</i> (<i>di-'æs-əm</i>)	'I look'

- b) In Rashti and the Rudbāri (Tatoid) group, both the Subjunctive and the Preterit have a *bV-* marker (see §6.5.6) but it only takes stress in the Subjunctive here. (See also §4.1, Pt. 10 on the difference in stress placement of WG vs. EG).
- c) Some languages have no *bV-* but still stress the initial syllable of the Subjunctive on the present root: Razajerdi *me-ruš-ém* > (Subj.) *rúš-em* 'sell'. Keringāni and Harzani have no *be-*, but stress is not indicated in the descriptions.

(82) Razajerdi

öštö=re ye hærf pórs-öm.
you-ABL one word ask:SBJ-1S₁
'Let me ask you something.'

Harzani
en öv=e gæræk riy-en.
this water=RA must pour-1S₁
'I must pour this water.' (MH: 483)

- d) Similarly to point (c), some languages delete the *bV-* morpheme before initial *v-* of the root, in which case the stress moves to the first syllable of the root, e. g., Nowkiāni *vénder-em* ~ *bé-vender-em* '(that) I stand/stay' (YSFN).
- e) If the verb root begins with a vowel, the vowel of *bV-* is lost, leaving the vowel of the root to bear the stress: Vafsi *b-ór-om* 'that I eat', Talyshi *b-árišt-o-m* '(that) I break', Lāhijani *b-ár-əm* 'that I bring'.

Additional Comments:

- 1) Some languages have small differences in Present tense vs. Subjunctive PAMs, Rashti 3SG. *xuré/búxurə* 'eats'; Kelārdashti 2SG. *bærné/béberi* 'you carry off'.

- 2) The labial of *bæ-* in Vafsi is often dropped or converted to *w-* or *y-*, depending on the preceding vowel: *ærgo wéssom* ‘I must go’, *tæ ðessey* ‘you should go’.
- 3) The Negative Imperative (or Prohibitive) takes the dedicated negator *mæ-* in Tatic, while Caspian languages have the usual *næ-* of the other tenses.
- 4) In only a few Tatic languages, the negator of the Subjunctive is *mæ-* as in the Prohibitive: *Nowkiāni čéra mæ-beræm-em?* ‘Why shouldn’t I cry?’ (YSFN), *Māsulei* (alternate forms) *mæ-dæ-færəs-əm ~ né-dæ-færəs-əm* ‘(that) I not ask’.
- 5) The negators *né-* and *mæ-* suppress the Subjunctive/Imperative *bæ-* prefix.
- 6) The singular Imperative has a Ø PAM; the plural is identical to the Subjunctive.
- 7) In S. Tati, the original Optative (§6.5.3) replaces the Subjunctive form only in the 2SG, e. g., *Chāli b-ášin-aš < b-ášin-a-š* ‘(that) you throw’.
- 8) Keringāni and Kalāsuri have a special situation for *transitive* Imperatives: in addition to the *bV-* marker, they optionally add the 3SG. Set₂ PAM, bleached of its usual marking of the Patient: Keringāni/Kalāsuri: *biždüž/bišdüüt* ‘sew!’, *bižvin/bišvin* ‘see!’, *bižgi/bišgin* ‘take!’ (all: YSFN).

6.5.2.1. Subjunctive paradigms

Lāhijāni:		No Preverbs		+ Preverbs	
1SG	<i>bá-bær-əm</i>	<i>bú-šu-m</i>	<i>dá-vəd-əm</i>	<i>hæ-de-m</i>	<i>ví-gir-əm</i>
2SG	<i>bá-bær-i</i>	<i>bá-š-i</i>	<i>dá-vəd-i</i>	<i>hæ-d-i</i>	<i>ví-gir-i</i>
3SG	<i>bá-bær-ə</i>	<i>bú-šu-n</i>	<i>dá-vəd-ə</i>	<i>hæ-de-Ø</i>	<i>ví-gir-ə</i>
1NEG	<i>ná-bær-əm</i>	<i>nú-šu-m</i>	<i>dá-n-vəd-əm</i>	<i>hæ-n-de-m</i>	<i>vín-gir-əm</i>
IMP.SG	<i>bá-bær</i>	<i>bú-šu</i>	<i>dá-vəd</i>	<i>hæ-di</i>	<i>ví-gir</i>
IMP.SG.NEG	<i>ná-bær</i>	<i>nú-šu</i>	<i>dá-n-vəd</i>	<i>hæ-n-di</i>	<i>ví-n-gir</i>
IMP.PL	<i>bá-bær-in</i>	<i>bá-š-in</i>	<i>dá-vəd-in</i>	<i>hæ-d-in</i>	<i>ví-gir-in</i>
IMP.PL.NEG	<i>ná-bær-in</i>	<i>ná-š-in</i>	<i>dá-n-vəd-in</i>	<i>hæ-n-d-in</i>	<i>ví-n-gir-in</i>
	‘take away’	‘go’	‘close, tie’	‘give’	‘pick up’
Rāmsari					
1SG	<i>bá-bor-om</i>	<i>bé-ber-om</i>	<i>bé-ber-əem</i>	<i>bé-ver-em</i>	<i>há-d-em</i>
2SG	<i>bá-bor-i/-e</i>	<i>bé-ber-i</i>	<i>bé-ber-i</i>	<i>bé-ver-i</i>	<i>há-d-i</i>
3SG	<i>bá-bor-ə</i>	<i>bé-ber-e</i>	<i>bé-ber-e</i>	<i>bé-ver-e</i>	<i>há-d-e</i>
1NEG	<i>ná-bor-om</i>	<i>né-ber-om</i>	<i>né-ber-əem</i>	<i>né-ver-em</i>	<i>næ-d-em</i>
IMP.SG/PL	<i>bá-bor(-in)</i>	<i>bé-ber(-in)</i>	<i>bé-ber(-in)</i>	<i>bé-ver(-in)</i>	<i>há-de/-d-in</i>
IMP.NEG	<i>ná-bor(-in)</i>	<i>né-ber(-in)</i>	<i>né-ber(-in)</i>	<i>né-ver(-in)</i>	<i>næ-de/-d-in</i>
	⇒	‘take away’	↔	‘give’	

	Chali	Leriki:	No Preverbs	+ Preverbs
1SG	<i>b-ášin-om</i>	<i>bí-kišt-ím</i>	<i>bí-dæ-m</i>	<i>pé-gæt-ím</i>
2SG	<i>b-ášin-aš (<-a-š)</i>	<i>bí-kišt-iš</i>	<i>bí-dæ-š</i>	<i>pé-gæt-iš</i>
3SG	<i>b-ášin-i</i>	<i>bí-kišt-í</i>	<i>bí-dæ-Ø</i>	<i>pé-gæt-í</i>
1NEG		<i>ní-kišt-ím</i>	<i>ní-dæ-m</i>	<i>pe-ní-gæt-ím</i>
IMP.SG/PL	<i>b-ášin/b-ášin-a</i>	<i>bí-kišt(-æn)</i>	<i>bí-dæ(-n)</i>	<i>pé-gi ~</i> <i>pé-gæt(-æn)</i>
IMP.NEG	<i>m-ášin/m-ášin-a</i>	<i>mæ-kišt(-æn)</i>	<i>mæ-dæ(-n)</i>	<i>pe-mæ-gi ~</i> <i>pemægæt(-æn)</i> <i>'pick up'</i>
	'throw'	'kill'	'give'	
Vafsi:	Cons.	-V	Consonant	-V
1SG	<i>bæ-vær-om</i>	<i>bæ-sso-m</i>	<i>dæ-gen-om</i>	<i>há-d-om</i>
2SG	<i>bæ-vær-i</i>	<i>bæ-sse-y</i>	<i>dæ-gen-i</i>	<i>há-de-y</i>
3SG	<i>bæ-vær-e</i>	<i>bæ-sso-(æ)</i>	<i>dæ-gen-e</i>	<i>há-d-o(æ)</i>
1NEG	<i>næ-vær-om</i>	<i>næ-sso-m</i>	<i>næ-gen-om</i>	<i>næ-d-om</i>
IMP.SG	<i>bæ-væ-Ø</i>	<i>bæ-ssæ-Ø</i>	<i>dæ-gen-Ø</i>	<i>há-dæ-Ø</i>
IMP.SG.NEG	<i>mæ-væ-Ø</i>	<i>mæ-ssæ-Ø</i>	<i>mæ-gen-Ø</i>	<i>mæ-dæ-Ø</i>
IMP.PL	<i>bæ-vær-a</i>	<i>bæ-ss-a</i>	<i>dæ-gen-a</i>	<i>há-d-a</i>
IMP.PL.NEG	<i>mæ-vær-a</i>	<i>mæ-ss-a</i>	<i>mæ-gen-a</i>	<i>mæ-d-a</i>
	'take away'	'go'	'fall'	'give'
				'pick up'

Those roots in Vafsi that palatalize their initial consonant after the Present TAM *æt-* revert to the original present root in the Subjunctive forms:

Root:	<i>-saz-</i>	<i>-ssæ-</i>	<i>(-šur-)</i>	<i>-zan-</i>
Palatalization:	<i>æt-s- > æč-</i>	<i>æt-s- > æč-</i>	<i>æt-š- > æč-</i>	<i>æd-z- > æj-</i>
Pres. 1SG	<i>æ-čaz-óm</i>	<i>æ-č-om</i>	<i>æ-čur-óm</i>	<i>æ(j)-jan-óm</i>
Subj. 1SG	<i>bæ-saz-om</i>	<i>bæ-ssø-m</i>	<i>bæ-šur-om</i>	<i>bæ-zan-óm</i>
Subj. 1SG.NEG	<i>næ-saz-om</i>	see	<i>næ-šur-om</i>	<i>næ-zan-om</i>
Imp. SG/(PL)	<i>bæ-saz(-a)</i>	previous	<i>bæ-šur(-a)</i>	<i>bæ-zan(-a)</i>
Imp. SG/(PL).NEG	<i>mæ-saz(-a)</i>	Table	<i>mæ-šur(-a)</i>	<i>mæ-zan(-a)</i>
	'build'	'go'	'wash'	'know'

Irregular Imperative forms

Root	<i>-a-</i>	<i>-ar-</i>	<i>-(h)or-</i>
Subj. 1SG	<i>b-a-ym</i>	<i>b-ar-óm</i>	<i>b-or-óm</i>
Imp. SG/PL	<i>b-úri/b-úri-a</i>	<i>b-óran/b-óran-a</i>	<i>b-óæ/b-ór-a</i>
Imp. SG.NEG	<i>m-úri/m-úri-a</i>	<i>m-óran/m-óran-a</i>	<i>m-óæ/m-ór-a</i>
	'come'	'bring'	'eat'

6.5.2.2. Uses of the subjunctive

Only a few representative uses of the Subjunctive can be mentioned in this chapter:

- 1) in subordinate verbs after modals (§6.6) and main verbs with control features;
- 2) in the protasis of conditional sentences implying doubt or uncertainty (§8.2.2);
- 3) in Purposive clauses (see §8.3.2);
- 4) in Optative senses (blessings and curses) if an Optative form is lacking (§6.5.3);
- 5) in the verb in an indefinite relative clause, as in the following two examples:

(83) Lāhijāni

hiš k̡s=a šənas-ən-i təmóm-ə i ko-n=a
 none person=RA know-PR-2S all-EZ this mountain-PL=RA
 bú-don-ə?
 PU-know-3s

‘Do you know anybody (that) knows all these mountains?’ (NJ: 127)

(84) Vafsi

esdæ qælém=i hær kay yə-wæ-Ø,, aer-go
 your pen=2s₂ each where PU-be-3s₁ DUR-want
 há-wend-uæ-Ø.
 PVB-find-PASS-3s₁

‘Your pen, wherever it is, has to be found.’

6.5.3. Optative

The Optative, lacking in Caspian, is quite robust in Tatic. It is formed by inserting the marker *-a-* into the Subjunctive after the Present stem. It expresses a wish or hope, as in blessings, curses, etc. The Optative form (Yarshater: Conditional Present) in Asālemi, Koluri, Gandomābi, Karani, and Kajali also has many Subjunctive uses (although there is an independent Subjunctive), esp. in conditional clauses (exs. 88–90). In Northern Talyshi, the Optative marker is raised to *o*. Within Tatic, the Optative is lacking in Vafsi, Khoini, the Tāromi group, and Māsulei, among others.

Forms of the Optative

	Leriki (N. Talyshi)	Asālemi (C. Talyshi)	Gandomābi (Central Tati)	Karani (Central Tati)	Koluri (Central Tati)
1SG	bí-hand-o-m	bá-š-a-m	be-dæv-a-ym	be-kær-a-ym	bé-čæk-a-m
2SG	bí-hand-o-ž	bá-š-a-š	be-dæv-a-š	be-kær-a-y	bé-čæk-a-š
3SG	bí-hand-o-Ø	bá-š-a-Ø	be-dæv-a-Ø	be-kær-a-Ø	bé-čæk-a-Ø
	‘read’	‘go’	‘run’	‘do’	‘break (tr.)’

Purely Optative sense:

- (85) *Leříki* Eštěhárdi (ST: 224)
hæmišæ b-o-ž! *xoda ji bé-köš-a-Ø*
 always be-OPT-2S₁ God he:OB PU-kill-OPT-3S₁
 ‘May you always be!’ ‘May God kill him!’

Karani (YSFN)

- yæman=er* *æ-ger-a-Ø.*
anthrax=2s₂ PVB-take-OPT-3s₁
 ‘May anthrax take you.’

In Leriki the Optative usually replaces the Subjunctive in the usual Subjunctive uses, even though both paradigms exist in the dialect. The Subjunctive and the Optative generally remain distinct in other Northern Talyshi dialects of Azerbaijan.

- (86) Leriki
 aey pi-dæ=ž-e hežo bi'-niňt-o-Ø kitob
 he:OB want:INF-LOC-3S₂AUX always SBJ-SIT-OPT-3S₁ book
 bá-hand-o-Ø.
 SBJ-read-OPT-3S₁
 'S/he always wants to sit and read books.'

(87) ovči ni'-be-Ø ki dí tifang-í aey bi'-žæn-o-Ø
 hunter NEG-be:PT-3S₁ SUB with rifle-OB he:OB SBJ-hit-OPT-3S₁
 'Wasn't there a hunter to shoot it with a gun?'

The original Optative may be used more commonly in the Present Conditional usage in Present/Future conditional clauses since the latter situations are more frequent:

- | | | |
|------|------------------------------------|-------------------------------|
| (88) | Asālemi | Gandomābi |
| | <i>ægær iya bá-b-a-š,...</i> | <i>ægæ bé-dæv-a-š,...</i> |
| | if here SBJ-be-OPT-2S ₁ | if PU-run-OPT-2S ₁ |
| | 'If you should be here, ...' | 'If you run ...' |

Many languages use the Optative in both senses:

- (89) Koluri
 Optative sense
 $xoda \quad xær-e \quad šax \quad mæ-da-ya-\emptyset!$
 God donkey-OB horn PROH-give-OPT-3S₁
 'May God not give horns to a donkey!' (AF: 428)

Conditional sense
 $\ægæ \quad dir \quad á-kær-a-nd,,\dots$
 if late PVB-do-OPT-3P₂
 'If they are late, ...'

(90) Karani (YSFN)

<i>xoda ešte pi-ær</i>	<i>bí-amorz-a-Ø!</i>
God your father-OB	PU-take-OPT-3S ₁
'May God absolve your father!'	
<i>aegæ bé-vrij-a-š,,</i>	<i>Ø-gen-i.</i>
if PU-run-OPT-2S ₁	PR-fall-2S ₁
'If you run, you will fall.'	

The Subjunctive is used in these senses in the languages that do not have an Optative:

(91) Māsulei

<i>xəda tə læ:næt bə-kær-ə!</i>	<i>Vafsi</i>
God you curse SBJ-do-3S ₁	God son-2S ₂ hold PU-have-3S ₁
'May God damn you!' (LM: 54, IV, 13)	'May God preserve your son!'

6.5.4. Future tense

While Formal Future tenses occur in some Caspian varieties, they are not often used. In Northern Tati and Northern Talyshi, the Future is quite robust. In Nowk-iāni and Alviri it is somewhat less regularly used and is very rare in Vafsi.

The Future formations consist of: A) an invariable form of 'want' plus Subjunctive of the verb; B) a finite form of 'want' plus an invariable form of the verb; C) the same Allative formation as in the Allative Present (see §6.5.1.3); or D) a Future participial form in *-æni(n)* (see next section also). Keringāni seems to have a unique Future formation (AAK 1954: 97–98): E) Infinitive (short form) + *-i-* ~ *-y-* + Copula (ex. 94).

My Lāhijāni fieldwork shows a Type A Future based on *xa-n* (Present of 'want' with a TAM but no PAMs). While Jahangiri (2003) does not discuss a formal Future, his examples show a similar Future without the TAM, *xa*, which also means 'must'.

	Type A	Type B	Type C	Type D
	Lāhijāni	Rashti	Leriki	Kalāsuri
1SG	<i>xa(-n) bU-šu-m</i>	<i>xa-yəm šo-on</i>	<i>b-æ-vot-é=m</i>	<i>šer-æni=m</i>
	'go'	'go'	'say'	'go'
	(Type C + D?)			
	Alviri			
	<i>be-vat-ænin-imæ</i>			
	'say'			

(92) Rashti

ha rā=ya xa-ídi šo-on.
 this.very road=RA want-3P go-INF
 ‘They will take (lit: go) this same road’ (RR: 240, 28)

Lāhijāni

guya xa bərf b-a-Ø.
 supposedly want snow PU-come-3s
 ‘It seems it is going to rain.’ (NJ: 170)

(93) Kalāsuri

em ešte pūl-i hard-æni=æ.
 this your money-OB eat-FUT=3s₁
 ‘He will embezzle (lit: eat) your money.’ (KL: 279)

Alviri

mella dærs-e nu be-vat-ænin-i.
 mullah lesson-EZ new ALL-say-FUT-3s₂
 ‘The teacher will teach (lit: say) a new lesson’ (HA: 173)

(94) Keringāni (Type E) (YSFN)

*te sæba če kur kard-e-y-šæ? mun sæba
 you tomorrow what work do-INF-FUT-2s₁ I tomorrow
 ši-e-y-ne šækar.
 go-INF-FUT-1s₁ hunt
 ‘What will you do tomorrow?’
 ‘I’ll go hunting tomorrow.’*

For examples of the Leriki Future, see exs. (45), (64), (237), (239), (270), (281), (317).

6.5.5. Future necessitative

In Leriki, which already has a regular Future tense, the participial morpheme *-ænin* forms a Future of Necessity, e. g., *hærd-ænín=im* ‘I will have to eat’. It also conveys a sense of “intention, strong desire to perform a specific action” (Miller 1953: 159).

(95) Leriki

az b-æy i dī kælimæ vot-ænin=im.
 I to-he:OB 1 2 word say-NEC=1s₁
 ‘I will have to tell him one or two words (i. e., about the problem)’

6.5.6. Past tense (Preterit)

- a. Most languages of the area use the same punctual *bV-* prefix as that of the Subjunctive as a marker of the Preterit (Intransitive and Transitive).
 - i. N. Tati, the enclaves Khoini and Dikin-Marāqei, and all Talyshi (except for the Jowkandāni area) lack the *bV-* punctual marker in the Preterit. For these languages, this marker will henceforth below be labeled as Subjunctive (SBJ), rather than punctual (PU), since this is its sole function here.
 - ii. Preverbs and the negator suppress *bV-* in the Preterit (as in Subjunctive).
 - iii. The Preterit *bV-* is generally stressed, except for WG, Rudbāri, and some neighboring Tati varieties, e. g., Koluri, Kafteji, Kelāsi, Nowkiāni.
- b. Some languages have a Ø PAM for the 3SG. in the Preterit, while others have a morpheme that is identical or similar to the 3SG. PAM of the Present.
- c. Most of Tatic makes a sharp distinction in intransitive and transitive conjugations in the Past system: Set₂ (“Oblique”) PAMs encode the transitive Agent while Set₁ (“Direct”) PAMs encode the intransitive S. Some languages additionally show the Absolutive nature of the Patient by cross-indexing it in the verb with Set₁ PAMs, but in general “there has always been a tendency to abandon Object-agreement in favour of A-agreement” (Haig 2008: 128).
- d. Caspian, Tatoid, and Razajerdi only have one (Intransitive type) conjugation.
- e. Set₂ PAMs encoding Past A are leftwardly mobile in Tatic (see §7.4.1).

6.5.6.1. Languages that take the punctual morpheme *bV-* in the preterit

Lāhijāni:

		+ Preverbs			
1SG	<i>bó-xord-əm</i>	<i>bú-šo-m</i>	<i>dá-kət-əm</i>	<i>há-da-m</i>	<i>ví-t-əm</i> < *ví-(g)it-əm
2SG	<i>bó-xord-i</i>	<i>bú-šo-i</i>	<i>dá-kət-i</i>	<i>há-da-y</i>	<i>ví-t-i</i>
3SG	<i>bó-xord-ə</i>	<i>bú-šo-Ø</i>	<i>dá-kət-ə</i>	<i>há-da-Ø</i>	<i>ví-t-ə</i>

‘eat’ ‘go’ ‘fall’ ‘give’ ‘pick up’

Rashti

		+ Preverbs	Tonekāboni	+ Preverbs	Sāravi
1SG	<i>ba-káft-əm</i>	<i>vi-ríšt-əm</i>	<i>bé-git-om</i>	<i>vé-git-om</i>	<i>bæ-merd-eme</i>
2SG	<i>ba-káft-i</i>	<i>vi-ríšt-i</i>	<i>bé-git-i</i>	<i>vé-git-i</i>	<i>bæ-merd-i</i>
3SG	<i>ba-káft-ə</i>	<i>vi-ríšt-ə</i>	<i>bé-git-e</i>	<i>vé-git-e</i>	<i>bæ-merd-e</i>

‘fall’ ‘get up, rise’ ‘get, catch’ ‘pick up’ ‘die’

Sāravi + Preverbs

dé-væss-eme

dé-væss-i

dé-væss-e

‘close, tie’

Vafsi

	Intransitive	Preverb	Transitive	Preverb
1SG	<i>bá-ræsa-ym</i> ‘arrive’	<i>dé-kætte-ym</i> ‘fall’	<i>b=ím-vattæ</i> ‘say’	<i>har=om-pærsa</i>
2SG	<i>bá-ræsa-y</i>	<i>dé-kætte-y</i>	<i>b=i-vattæ</i>	<i>har=i-pærsa</i>
3SG	<i>bá-ræsa-Ø</i>	<i>dé-kættæ-Ø</i>	<i>b=ís-vattæ</i>	<i>har=es-pærsa</i>
NEG	<i>né-ræsa-ym</i>	<i>ná-kætte-ym</i>	<i>n=ím-vattæ</i>	<i>n=ím-pærsa</i>

Central Tati

N. Talyshi
(transitional)

	Nowkiāni	transitive	Kafteji	Koluri	Jowkandāni
1SG	<i>be-xót-im</i>	<i>be-váét-em</i>	<i>bá-káét-im</i>	<i>be-váéšt-im</i>	<i>bá-ræs-ím</i>
2SG	<i>be-xót-iš</i>	<i>be-váét=er</i>	<i>bá-káét-iš</i>	<i>be-váéšt-iš</i>	<i>bá-ræs-iš</i>
3SG	<i>be-xót-Ø</i>	<i>be-váét=eš</i>	<i>bá-káét-Ø</i>	<i>be-váéšt-Ø</i>	<i>bá-ræs-i=e</i>
3F	<i>be-xót-æ</i> ‘sleep’	<i>be-váét-æ</i> ‘weave’	<i>bá-káét-æ</i> ‘fall’	<i>bá-káét-æ</i> ‘jump’	<i>bá-káét-æ</i> ‘arrive’

Tatoid: Tutkāboni (Rudbāri)

Gurāni (Tāleqāni)

	Intrans. Past	Trans. Past	Intrans. Past	Trans. Past	Past
1s	<i>bá-mórd-om</i>	<i>bá-páét-om</i>	<i>bá-xot-om</i>	<i>bá-xord-om</i>	<i>há-git-om</i>
2s	<i>bá-mórd-i</i>	<i>bá-páét-i</i>	<i>bá-xot-i</i>	<i>bá-xord-i</i>	<i>há-git-i</i>
3s	<i>bá-mórd-Ø</i>	<i>bá-páét-Ø</i>	<i>bá-xot-Ø</i>	<i>bá-xord-Ø</i>	<i>há-git-Ø</i>

6.5.6.2. Languages that do not take the Punctual morpheme *bV-* in the preterit

	Khoini	transitive	Harzani	transitive	Kalāsuri	transitive
1SG	<i>genést-im</i>	<i>vínd=em=(e)</i>	<i>šér-in</i>	<i>öt=ma</i>	<i>genést-im</i>	<i>mæs=em=æ</i>
2SG	<i>genést-iš</i>	<i>vínd=i=(e)</i>	<i>šér-i</i>	<i>öt=la</i>	<i>genést-iš</i>	<i>mæs=er=æ</i>
3SG	<i>genést-e</i> ‘fall’	<i>vínd=eš=(e)</i> ‘see’	<i>šér-æ</i> ‘go’	<i>öt=ja</i> ‘say’	<i>genést-æ</i> ‘fall’	<i>mæs=eš=æ</i> ‘hear’

Leriki

	intransitive		transitive	
1SG	<i>omáé-ym</i> ‘come’	<i>vít-im</i> ‘run’	<i>gíní-m</i> ‘fall’	<i>kárd=ím=e</i> ‘do’
2SG	<i>omáé-yš~ž</i>	<i>vít-iš~ž</i>	<i>gíní-š~ž</i>	<i>kárd=í=e*</i> <i>kárd=íš=e ~</i>
3SG	<i>omáé-y</i>	<i>vít-e</i>	<i>gíní-Ø</i>	<i>kárd=íš=e</i> <i>kárd=íž=e</i> ‘pick up’

*These forms are generally avoided for phonological reasons.

Dikin Marāqeи, in addition to the regular intransitive and transitive conjugations, has an intransitive-like past Objective conjugation. In Razajerdi only a few verbs ('run, buy, hear, slice, read, break') take the punctual marker, while others do not.

Dikin Marāqeи			
	intransitive	transitive	Objective
1SG	<i>genést-in</i>	<i>vínd=em=æ</i>	<i>vínd-in</i>
2SG	<i>genést-iš</i>	<i>vínd=et=æ</i>	<i>vínd-iš</i>
3SG	<i>genést-æ</i>	<i>vínd=eš=æ</i>	<i>vínd-æ</i>
3F	<i>genést-ian</i>		<i>vínd-ian</i>
	'fall'	'see'	'X saw me, you, etc.'

Razajerdi			
(intransitive/transitive)			
<i>kæt-öm</i>	<i>göt-öm</i>	<i>bə-xúnd-em</i>	
<i>kæt-i</i>	<i>göt-i</i>	<i>bə-xúnd-i</i>	
<i>kæt-Ø</i>	<i>göt-Ø</i>	<i>bə-xúnd-Ø</i>	
<i>kæt-æ</i>	<i>göt-æ</i>	<i>bə-xúnd-æ</i>	
'fall'	'say'	'read'	

6.5.7. Imperfect tense

The various types of durative and/or Imperfect markers in Caspian and Tatic are:

- A) Prefixes or Ø: 1) Ø; 2) *æt-* durative; 3) *me-* ~ *eN-* durative; 4) Allative markers; 5) the so-called Augment *æ-* found only in Northern and Central Talyshi and only in combination with the *-i-* suffix of B3 below. Languages with prefixes A2–4 share these prefixes with the Present tense.
- B) Suffixes (not shared by the Present): 1) *-enV-*; 2) *imi* (Kelärdashti only); 3) *i* ~ *æy* (along with the Augment in Northern/Central Talyshi, but no Augment in Southern Talyshi, the neighboring Tati dialects, or W. Gilaki).

Comments:

- 1) The Ø Imperfect marker based on the Past stem is found across the whole Caspian family, excluding WG and Kelärdashti. These languages generally do not coincide with those that have a Ø marker in the Present tense.
- 2) When a verb has a lexical preverb in languages with a Ø Imperfect marker, the Preterit and the Imperfect are indistinguishable since the preverb suppresses the punctual *bV*: Lähijāni: *há-da-m* 'I gave; used to give'.
- 3) In Central Caspian (except for Kelärdashti) there are two Imperfect types, Ø and *enV*; the latter has both irrealis and habitual or even progressive senses but in EG it is exclusively a conditional marker (see §6.5.11–§6.5.12).
- 4) WG is the only Caspian variety that has the *-i-* Imperfect marker; WG also lacks the conditional marker *-enV-*.

- 5) The *-i-* marker is also used with the past AUX (verb + *b-i-Ø*, etc.) in irrealis paradigms (§6.5.11–§6.5.12).
- 6) The Imperfect in some languages is formed on the Present stem: Talyshi, Koluri, Kelārdashti (*-imi-*), and partly *-ena-* in Rāmsari and Tonekāboni.
- 7) The *aey* suffix of some C. Tati (and mobile clitic of Keringāni Past Subjunctive, §6.5.11) is the same as *-i-* of Points 4–5: *aey* > *ē* > *i*. It may derive from *hait*, the Old Iranian 3rd sg. optative of ‘be’ (Windfuhr 1987: 393).

6.5.7.1. Ø-marked imperfects

	Lāhijāni I	final -V	Preverb	Rāmsari I	Tonekāboni I
1SG	Ø-xónd-əm	Ø-bye-m	há-da-m	Ø-gút-əm	Ø-gút-əm
2SG	Ø-xónd-i	Ø-b-i	há-da-y	Ø-gút-e	Ø-gút-e
3SG	Ø-xónd-ə	Ø-be-Ø	há-da-Ø	Ø-gút-ə	Ø-gút-ə
	‘read’	‘cut’	‘give’	‘say’	‘say’
Sāravi					
			(Preverb suppressed in durative tenses)		
1SG	Ø-merd-eme	Ø-ši-me	Ø-væss-eme	Ø-get-eme	Ø-viness-eme
2SG	Ø-merd-i	Ø-ši-i	Ø-væss-i	Ø-get-i	Ø-viness-i
3SG	Ø-merd-e	Ø-ši-ye	Ø-væss-e	Ø-get-e	Ø-viness-e
	‘die’	‘go’	‘tie, close’	‘take, catch’	‘see’

6.5.7.2. Durative markers shared with present: *aet-* ~ *aer-* and *me-* ~ *eN-*

	Vafsi	transitive	Hezārrudi	Eshtehārdi	Khoini
1SG	ær-kætte-ym	im-ær-vattæ	me-kæt-im	mi-væšt-im(ae)	"ut-m-išt-im
2SG	ær-kætte-ey	i-r-vattæ	me-kæt-i	mi-væšt-iš(ae)	"ut-m-išt-iš
3SG	ær-kætte-Ø	is-ær-vattæ	me-kæt-Ø	mi-væšt-Ø	"ut-m-išt-e
	‘fall’	‘say’	‘fall’	‘run’	‘rise, get up’

6.5.7.3. Imperfect *-i-* ~ *-aey-* formant, added to past stem

	Caspian	Rudbāri	Central Tati	Karnaqi	Gandomābi
1SG	Rashti (WG)	Tutkāboni	Dizi		
	gúft-i-m	gut-i-m	xærd-ey=m=e	=em xand-ey	kætt-aey-m
2SG	gúft-i (< -i-i)	gut-i (< -i-i)			kætt-aey-š
3SG	gúft-i-Ø	gut-i-Ø			kætt-aey-Ø
	‘say’	‘say’	‘eat’	‘read’	‘fall’

This formation in Keringāni is only used as a Past Subjunctive (§6.5.11).

Imperfect formants *-i-* ~ *-æy-* and *-imi-*, added to Present stem

	Languages with Augment + <i>-i-</i>		<i>-i-</i> , no Augment	
	Leriki	Asālemi	Māsulei	Koluri
PR	<i>žæn-</i>	<i>hæ-</i>	<i>virij-</i>	<i>pe-</i>
PT	<i>žæ-</i>	<i>hard-</i>	<i>virit-</i>	<i>pætt-</i>
1SG	<i>æ-žæn-i-m</i>	<i>æ-hæ-y-m</i>	<i>æ-vrij-i-m</i>	<i>p-i-m</i>
2SG	<i>æ-žæn-i-š</i>	<i>æ-hæ-y-š</i>	<i>æ-vrij-i-š</i>	<i>p-i-š</i>
3SG	<i>æ-žæn-i-Ø</i>	<i>æ-hæ-y-Ø</i>	<i>æ-vrij-i-Ø</i>	<i>p-i-Ø</i>
1PL	<i>æ-žæn-i-mun</i>	<i>æ-hæ-y-mon</i>	<i>æ-vrij-i-mun</i>	<i>p-i-mon</i>
2PL	<i>æ-žæn-i-n</i>	<i>æ-hæ-y-n</i>	<i>æ-vrij-i-run</i>	<i>p-i-n</i>
3PL	<i>æ-žæn-i-n</i>	<i>æ-hæ-y-n</i>	<i>æ-vrij-i-n</i>	<i>p-i-n</i>
	‘hit, strike’	‘eat’	‘run’	‘cook’
				‘say’

-imi- Suffix

Kelārdašti

*el-**ešt-**el-imi-yæme**el-imi-Ø**el-imi-yæ**el-imi-yæmi**el-imi-yænne**el-imi-yænne*

‘put’ (KK: 81)

6.5.7.4. Allative-marked imperfects

This Imperfect is formed similarly to the Allative-type Present (*be-*) but with the past AUX. (Also see §6.5.8.1 for these forms used as Past Progressives in some areas.)

	Keringāni: ‘do’	Kalāsuri: ‘sleep’	Harzani: ‘say’
1SG	<i>be-kard-e bi-ne</i>	<i>be-het-é=u-m</i>	<i>be-öt-e ber-in</i>
2SG	<i>be-kard-e bi-še</i>	<i>be-het-é=u-š</i>	<i>be-öt-e ber-i</i>
3SG	<i>be-kard-e biæ-Ø</i>	<i>be-het-é=u-Ø</i>	<i>be-öt-e ber-æ</i>
ALL-VB:PT-INF		ALL-VB:PT-INF=	ALL-VB:PT-INF
AUX:PT-SET ₁		AUX:PT-SET ₁	AUX:PT-SET ₁

6.5.7.5. *Uses of the imperfect*

In addition to the usual past habitual sense, the Imperfect may have a progressive sense since the Progressive forms, where they exist, are not necessarily obligatory. The Imperfect also has Conditional and Past Subjunctive uses, but there are also dedicated forms for the latter paradigms.

- (96) Lāhijāni
yeblen gušt Ø-xord-əm. čiči Ø-gút-i.
 before meat IMF-eat:PT-1S what IMF-say:PT-2S
 ‘I used to eat meat before.’ ‘What were you saying?’
- (97) Sāravi
xæzune bar=e Ø-værd-e, Ø-rut-e.
 treasury load=RA IMF-take:PT-3S IMF-sell:PT-3S
 ‘He would take the load of the treasury (and) would sell (it).’ (SY: 68)
- (98) Leriki
texnikum=aedæ a-hand-i-mon.
 technikum-at DUR-read-IMF-1P₁
 ‘We studied at a technical school.’
- Vafsi
in dokan=dæ kár=es aer-kærd.
 this shop=at work=3S₂ DUR-do:PT
 ‘He used to work in this shop.’
- Conversely, Conditionals (§6.5.12) are often interchangeable with the Imperfect in habitual senses in Rāmsari and Tonekāboni but, so far as is known, not in EG.
- (99) Rāmsari
ruz xót-enæ=ba-n, šæb hæmlá Ø-körd-ən.
 day sleep:PT-COND=AUX:PT-3P night attack IMF-do:PT-3P
 ‘They would sleep in the daytime (and) would attack at night.’
- (100) Tonekāboni
berénj-e koléš=e vé-git-ene ba-n, súj-aene
 rice-REZ bran=RA PVB-take:PT-COND AUX:PT-3P VB-burn-COND
ba-n, vé=re ælæk Ø-górd-en.
 AUX:PT-3P it-RA sieve IMF-do:PT-3P
 ‘They used to take the rice bran, burn (it and) sift it.’ (Amini: 115)

6.5.8. Progressive aspect (present and past)

As outlined below, there are two main types of Progressives in the area, with various subtypes. Both may appear together in some languages, creating a mixed third type.

6.5.8.1. Type I: Locative Formations

Type IA: The Progressive is formed in N. Tati, Lerdi, Chāli, Khoini, and others according to the formula: INFINITIVE OR PAST STEM + LOCATIVE + COPULA.

In Type IA the Locative postposition *ændæ* of Harzani has three variant forms: 1) after a vowel the first vowel of the suffix is lost and with all three variants the last vowel is also lost before the vowel of the PAM: *ama-* > *ama-nd-en* ‘I am coming’; 2) after a root ending in *-t*, the *-ændæ-* suffix loses the first syllable: *öt-* > *öt-d-en* ‘I am saying’; 3) after fricatives and sonorants the *-n-* remains but the *-d-* is lost: *bəs-* > *bəs-n-en*.

In Type IB the locative formant of Kalāsuri-Khoynarudi is deleted leaving only the infinitival and the copular elements. They also have an unusual copula formed by a demonstrative (*en* ‘this’, *œv* ‘that’) + an enclitic copula. In the Progressive, the demonstrative attracts the copula leftwards. An overt Object usually suppresses the demonstrative and hosts the copula (101a), but both may appear together (101b).

Type IA: Harzani

	Pres. Progressive	Past Progressive
1SG	<i>öt-d-en</i>	<i>bəs-n-en</i>
2SG	<i>öt-d-e</i>	<i>bəs-n-e</i>
	VB-LOC=AUX ‘say’	vb-LOC AUX;PT-PAM ‘close’
		‘say’
		‘close’

Type IB: Kalāsuri

Pres. Prog.	Past Prog.
<i>en=em kust-e</i>	<i>en=u-m kust-e</i>
<i>en=iš kust-e</i>	<i>en=u-š kust-e</i>
DEM=PAM	VB-INF
‘pound’	

- (101) Kalāsuri (KL: 278)

<i>olæt=em</i>	<i>šušt-e.</i>	<i>œv=u=š</i>	<i>olæt šušt-e.</i>
clothes=1s ₂	wash-INF	that=AUX:PT=2s ₂	wash-INF
‘I am washing clothes.’		‘You were washing clothes.’	

While Leriki uses IA typology for the Present tense and has no independent Present Progressive, it does have a Past Progressive of this type (Pirejko: "Past Dative II"; Miller: "Compound Imperfect II") that formally contrasts with the Imperfect (§6.5.7.3), but is often interchangeable with it. Pirejko (1976: 337) states that this "...tense form denotes any durative action in the past."

	Full form -V	Full form -t/d	Short Form -t/d
	STEM+INFINITIVE+LOCATIVE+		STEM (stressed)+LOCATIVE+COPULA:PT
	COPULA:PT		
1SG	š-é-dæ=bi-m	vot-é-dæ=bi-m	vót-dæ=bi-m
2SG	om-é-dæ=bi-š	mand-é-dæ=bi-š	mánd-dæ=bi-š
	'go, come'	'say, stay'	'say, stay'
(102)	Leriki		
	ayil vit-dæ=b-e	bæ di	mašin-i.
	child run-LOC=AUX:PT-3S ₁	to	after car-OB
	'The child was running after the car.'		
	penj sor vaxt do-y-dæ=b-in.		
	five year time give-INF-LOC=AUX:PT-3P ₁		
	'They used to give five years leeway.'		

Type IC (only Caspian and Rudbāri): instead of a postposition, the copula is a Locative Be-verb (BE₄, §6.8). The infinitive is: a full form (Rashti); a short form, identical to the past stem (Tonekāboni); or the -n of the infinitive is lost (Lāhijāni):

Lāhijāni, Progressive I

	Present	Past		
1SG	xord-é dær-əm	xord-é də-bu-m		
	eat:PT-INF BE ₄ -1S	eat:PT-INF BE ₄ -AUX:PT-1S		
1NEG	xord-e dán-ny-em	xord-e dó-nye bu-m		
	'eat			
	Present	Past		
1SG	ha-dá dær-əm	ha-dá də-bu-m		
	PVB-give:PT BE ₄ -1S	PVB-give:PT BE ₄ -AUX:PT-1S		
1NEG	ha-da dán-ny-em	ha-da dó-nye bu-m		
	'give'			
Rashti Prog. I	Langerudi	Rāmsari	Tonekāboni	Tutkāboni
	Prog. I		Prog. I	(Rudbāri)
gust-án dær-əm	gut-é dær-əm	za dær-əm	xárd dár-om	xard-en dár-om
'say'	'say'	'hit'	'eat'	'eat'

(103) Lāhijāni

Present Progressive I

<i>mu</i>	<i>yəza</i>	<i>xord-é</i>	<i>dər-əm.</i>
I	food	eat-INF	BE ₄ -1S
'I am eating.'			

Past Progressive I

<i>hava</i>	<i>rošen=a</i>	<i>bo</i>	<i>dər-bu-Ø</i>
weather	bright=CS	become:INF	BE ₄ -AUX:PT-1S
'The weather was clearing up.'			

(104) Rāmsari (SR: 146)

Langerudi

<i>kemi</i>	<i>ra</i>	<i>ša</i>	<i>dær-e?</i>	<i>adæm</i>	<i>æmá</i>	<i>dær-e.</i>
which	road	go:INF	BE ₄ -3S	person	come:INF	BE ₄ -3S
'Which road are you taking?'						

Type ID: In Kelārdashti and Gurāni (Tatoid) fully conjugated Present and Past forms of the Locative Be-verb (Be₄) precede the Present and Imperfect:

Kelārdashti: 'to carry off, take away'

Present Progressive Past Progressive

1SG	<i>dær-eme</i>	<i>bær-eme</i>	<i>dæ-bi-yæme</i>	<i>bær-imyæme</i>
	BE ₄ -AUX:1S	take-1S	BE ₄ -AUX:PT-1S	take-IMF-1S

Gurāni: 'to close; bake'

Present Progressive

<i>der-im</i>	<i>dé-m-bænd-im</i>
BE ₄ -1P	PVB-DUR-bake-1P

Type IE: In Mazanderani and Dikin Marāqeи, an invariable 3SG of the Present/Past of Be₄ precedes the verb. Mazanderani may also optionally delete the PAM (*dær-e* > *dær*) or conjugate both verbs, probably under Persian influence.

Sāravi: 'come'

'eat'

Dikin-Marāqeи:

'become'

Pres. Progressive

Past Progressive

Pres. Progressive

1SG	<i>dær(-e)</i>	<i>e-m-be</i>
2SG	<i>dær(-e)</i>	<i>e-n-i</i>
3SG.M	<i>dær(-e)</i>	<i>e-n-e</i>
3SG.F		

<i>dæy(-ye)</i>	<i>Ø-xord-eme</i>
<i>dæy(-ye)</i>	<i>Ø-xord-i</i>
<i>dæy(-ye)</i>	<i>Ø-xord-e</i>

<i>dær-i</i>	<i>me-b-in</i>
<i>dær-i</i>	<i>me-b-i</i>
<i>dær-i</i>	<i>me-b-ik</i>
<i>dær-i</i>	<i>me-b-igan</i>

BE ₄ (-3S)
come-PR-PAM

BE ₄ :PT(-3S)
eat-IMF-PAM

BE ₄ 3S ₁
DUR-become-PAM

(105) Sāravi

bé-di-ye zena dær e-n-e.
 PU-SEE:PT woman BE₄:PR come-PR-3S
 'He saw that (his) wife is coming.' (SY: 66)

at-ta zena at-ta mærdi dæy ši-ne.
 1-UNC woman 1-UNC man BE₄:PT go:IMF-3P
 'A man and a woman were going (along).' (SH: 155)

(106) vešun dær-ene šu-n-ene sere.

they BE4-3P go-PR-3P house
 'They are going home.' (SH: 149)

dæy-ne æli=re šekænje Ø-kard-ene.
 BE₄:PT-3P P.N.=RA torture IMF-DO:PT-3P
 'They were torturing Ali.' (SY: 68)

6.5.8.2. Type II: Independent particles derived from *kăr* 'work'

In type IIA an invariable particle, *kVrV* (< *kăr* 'work, doing', Windfuhr 1989: 256), accompanies durative tenses to form a Progressive or convey a sense of Imminence ('be about to'). This type occurs in Rashti, Southern and Central Talyshi, and in adjacent varieties of Tati. Some languages use more than one of these strategies.

Invariable Particle *kVrV* + Main Verb

Type IIA: *kárə/kára/káræ/kora* precedes the regular Present or Imperfect tenses

	Mäsolei	Keläsi	Gandomäbi	Karani Prog. I
1SG	<i>kärə xəs-Ø-ám</i>	<i>kæræ me-xos-ám</i>	<i>kora xos-ém</i>	<i>kari xoss-ém</i>
2SG	<i>kärə xəs-Ø-é</i>	<i>kæræ me-xos-í</i>	<i>kora xos-í</i>	<i>kari xoss-í</i>
3SG	<i>kärə xəs-Ø-ð</i>	<i>kæræ me-xos-ð</i>	<i>kora xos-é</i>	<i>kari xoss-eæ</i>
3F				<i>kari xoss-iæ</i>
1PL	Koluri	Rashti Prog. II	Hezärrudi	Asälemi Prog. I
	<i>kera xes-Ø-ém</i>	<i>kára xus-Ø-ími</i>	<i>kæræ xos-ám</i>	<i>káræ bæ-xüt-í-mun</i>
2PL	<i>kera xes-Ø-é</i>	<i>kára xus-Ø-ídi</i>	<i>kæræ xos-á</i>	<i>káræ bæ-xüt-í-run</i>
3PL	<i>kera xes-Ø-énd</i>	<i>kára xus-Ø-ídi</i>	<i>kæræ xos-énde</i>	<i>káræ bæ-xüt-í-n</i>
	'sleep'			

Asälemi and Karani Progressive I forms above mostly indicate Imminence, while their Progressive II forms (types IIB, IID, resp.) are used as Progressives. The

Māsulei form indicates both senses. In Type IIB (Kajali) *ko(re)* follows Main Verb and in Type IIC (Kafteji) *kérda*, a locative form, precedes the verb.

	Type IIB	Type IIC
	Kajali	Kafteji
1SG	<i>me-vræm-em kore.</i> ‘I am crying/weeping’	<i>kérda me-š-ám.</i> ‘I am going’

Progressive *kar* particle combined with Be-verbs

IID: In Type IID the Asālemi Present Progressive II is formed with the particle *kar* preceding the verb in the infinitive form and a mobile copular element, usually encliticized to *kar* or even farther leftward. There is also a shortened form of *kar* >*ka*:

	Present	Short form	Past	Short form
	Progressive		Progressive	
1SG	<i>kár=im vat-è</i>	<i>ká=m vat-è</i>	<i>kár=b-im vat-è</i>	<i>ká=b-im vat-è</i> ‘say’

Charozh, the Central Talyshi enclave in the N. Talyshi zone (§2.2.1), has shifted this short-form progressive (with N. Talyshi vowel raising: *ka* > *ko*) to the sole Present tense. The Allative-marked C. Talyshi Present is the Charozh Future, as in N. Talyshi:

(107) Charozh

čo	sahat	dærs	<i>ko=mun</i>	<i>do-y.</i>
four	hour	lesson	PROG=1P ₁	give-INF
‘We teach for four hours.’				

tikæ	æmæ	<i>zü=mun</i>	<i>ko</i>	<i>š-e.</i>
bit	we	early=1P ₁	PROG	go-INF
‘We go a little early.’				

IIE: Lāhijāni Progressive II inserts the particle *ka* (< *kar*) between a short infinitive of a simplex verb and the Be₄. Type IIE is especially common with compound verbs formed with the light verb *kun/gud-* ‘do’, in which case *kun/gud-* is deleted and the cognate *ka* takes its place between the NVE and Be₄:

	IIC: Lāhijāni II (Simplex Verb)	
	Pres. Progressive	Past Progressive
1SG	<i>gut-è=ká=dər-əm</i>	<i>gut-è=ká=də=bu-m</i>

‘I am saying’

(Compound Verb with *gudən*)

Pres. Progressive

bazi gudən > *bazi ka=dər-əm*
 ‘to play’ ‘I am playing’

IIF: The *kar* element in Karani and Karnaqi consists of the full Present or Imperfect tense of the verb *kær-* ‘do’, following the invariable past participle of the main verb.

IIG: In Eštehārdi the fully conjugated Present/Imperfect form of the main verb is followed by *kar/kær* plus a conjugated form of *Be₄*:

	IID: Karani/Karnaqi Prog. II Pres. Progressive	IIE: Eštehārdi Pres. Progressive
1SG	<i>mærd-æ kær-em</i> ‘I am/was dying’ (YSFN)	<i>mi-væz-em kar/kær=dær-imæ</i> ‘I am running’ (ST: 225)

6.5.8.3. Type III: Simultaneous use of *Be₄* and the *kára* particle

	Rashti Prog. III
1SG	<i>kára guft-án=dær-əm</i> ‘I am saying’

Some languages, e. g., Vafsi, have no strategies for marking Progressives, which are subsumed under the regular Present and Imperfect tenses, as in exs. 194, 284, 285B, 336, and the first sentence of the Vafsi text (§10.4).

6.5.9. Present perfect tenses

In my analysis, Caspian varieties all lack a *Present Perfect* which has merged with the Preterit. Resultative forms (‘he is sitting/seated in the chair’) in Caspian languages should not be confused with a true dynamic Perfect (‘I have often sat in that chair’). Some grammars list Caspian Perfects formed with the auxiliary *dašt-* ‘have’, but I have never observed such forms in actual speech or in texts. Jahāngiri (2003) makes no mention of Present Perfect paradigms and he translates the Preterits of many Lāhijāni sentences using the Present Perfect in Persian.

The stress patterns in most Tatic varieties differ between the Preterit (stress on the past stem) and the Present Perfect (stress on the participial formant).

	Kelāsi	Kajali (YSFN)	Gandomābi (YSFN)
	Preterit Pres. Perf.	Preterit Pres. Perf.	Preterit Pres. Perf.
1SG	<i>bə-kədət-</i> <i>im</i> ‘fall’	<i>bə-kət-</i> <i>é-m</i> ‘fall’	<i>be-kætt-</i> <i>im</i> ‘fall’

	Preterit	Pres. Perf.	Preterit	Pres. Perf.	Preterit	Pres. Perf.
1SG	<i>be-bræmés-</i> <i>im</i> ‘cry’	<i>bræmes-</i> <i>é-m</i> ‘cry’				

	Chāli (ST:197)	Jubani (Rudbāri)	Khoini (YX:176)
	Preterit Pres. Perf.	Preterit Pres. Perf.	Preterit Pres. Perf.
1SG	<i>be-téét-im</i> ‘run’	<i>be-tæt-</i> <i>é-m</i> ‘say’	<i>bu-gút-</i> <i>om</i> ‘laugh’
	Razajerdi (YSFN)	Harzani (AAK: 116–18) Asālemi	
	Preterit Pres. Perf.	Preterit Pres. Perf.	Preterit Pres. Perf.
1SG	<i>káét-ay</i> (2 f.) ‘fall’	<i>kæt-áy</i> (2 f.) ‘fall’	<i>yér=</i> <i>m=æ</i> ‘hit’
	Māsulei:	Intransitive ‘burn’	Transitive ‘eat’
	Preterit Present Perfect	Preterit Present Perfect	
1SG	<i>sí-m=æ</i>	<i>sié=m=æ</i>	<i>háerd=əm=æ</i>
2SG	<i>sí-š=æ</i>	<i>sié=š=æ</i>	<i>háerd=i=æ</i>
3SG	<i>sí-Ø=æ</i>	<i>sié=Ø</i>	<i>háerd=əš=æ</i>

Some languages, in addition to stress differences, also have slightly different PAMs in the Preterit vs. the Present Perfect as in the Leriki Intransitive forms below:

Leriki: Preterit vs. Present Perfect

Stems ending in Vowel	
Pret.	Pres. Perf.
1SG	<i>omé-ym</i>
2SG	<i>omé-yž¹</i>
3SG	<i>omé-y</i>
1PL	<i>omé-ymon</i>
2PL	<i>omé-žon¹</i>
3PL	<i>omé-yn</i>
1NEG	<i>nómæym</i> ‘come’

Stems ending in Consonant

Preterit	Pres. Perf.	Preterit	Pres. Perf.
<i>vít-im</i>	<i>vit-áe-m</i>	<i>kárd=im=e</i>	<i>kard-áe=m=e</i>
<i>vít-iž¹</i>	<i>vit-áe-ž¹</i>	<i>kárd=e²</i>	used only fronted ²
<i>vít-e</i>	<i>vit-áe-Ø</i>	<i>kárd=iž=e¹</i>	<i>kard-áe=ž=e¹</i>
<i>vít-imon</i>	<i>vit-áe-mon</i>	<i>kárd=imón=e</i>	<i>kard-áe=mon=e</i>
<i>vít-ižon¹</i>	<i>vit-áe-žon¹</i>	<i>kárd=ín=e</i>	<i>kard-áe=n=e</i>
<i>vít-in</i>	<i>vit-áe-n</i>	<i>kárd=ižon=e¹</i>	<i>kard-áe=žon=e¹</i>
<i>nívitim</i>	<i>vité nim</i>	<i>níkárdíme</i>	<i>kardé nime</i>
‘run’		‘do, make’	

1) the *-ž* in all PAMs can always alternate with *-š*; 2) these forms are avoided for phonological reasons (*kárd-e* < *kárd=i=e*) but this is not a problem since the Set₂ PAMs are commonly fronted (§7.4.1);

(108) Leriki

<i>ve</i>	<i>say</i>	<i>be</i>	<i>ki</i>	<i>tí</i>	<i>omáé=ž.</i>
very	healthy	be	SUB	you	come:PT=2S ₁
'Thank you very much for having come (lit: that you have come).'					
<i>gæst=o=žon=e</i>			<i>vot-áé.</i>		
spite=from=3P2=AUX			say:PT-PPL		
'They have said (it) out of spite.'					

In a slightly different pattern, the stress in Vafsi and possibly Sagzābādi falls on the Punctual marker in the Preterit and on the *verb stem* in all Perfect forms. In Vafsi the Past Participle *-áé* is replaced in the Present Perfect by the Perfect formant *-e*, which is only apparent in the 3SG/PL of Intransitives and all persons of the Transitives. The *-e* formant is not used in other Perfect paradigms (§6.5.10). With past stems ending in *-a* in Vafsi, the stress is the only way to distinguish the Preterit (*bé-resa-ym*) from the Perfect (*bæ-resá-ym*), as in (109).

Vafsi: Preterit vs. Present Perfect

Intransitive		Transitive	
Preterit	Pres. Perf.	Preterit	Pres. Perf.
1SG <i>bé-mærd-e-ym</i>	<i>bæ-mærd-e-ym</i>	<i>b-ím-rutt-æ-Ø</i>	<i>b-im-rútt-e</i>
2SG <i>bé-mærd-e-y</i>	<i>bæ-mærd-e-y</i>	<i>b-i-rutt-æ-Ø</i>	<i>b-i-rútt-e</i>
3SG <i>bé-mærd-æ-Ø</i>	<i>bæ-mærd-e</i>	<i>b-is-rutt-æ-Ø</i>	<i>b-is-rútt-e</i>
1PL <i>bé-mærd-e-yam</i>	<i>bæ-mærd-e-yam</i>	<i>b-óan-rutt-æ-Ø</i>	<i>b-oan-rútt-e</i>
2PL <i>bé-mærd-e-ya</i>	<i>bæ-mærd-e-ya</i>	<i>b-ian-rutt-æ-Ø</i>	<i>b-ian-rútt-e</i>
3PL <i>bé-mærd-æ-nde</i>	<i>bæ-mærd-e-nde</i>	<i>b-isan-rutt-æ-Ø</i>	<i>b-isan-rútt-e</i>
1NEG	<i>né-mærde-ym</i>	<i>n-ím-rutt-æ-Ø</i>	<i>n-ím-rutt-e</i>
3NEG	<i>né-mærdæ-Ø</i> 'die'	<i>n-is-rutt-æ-Ø</i>	<i>n-is-rutt-e</i> 'sell'

(109) Vafsi

<i>yawa</i>	<i>soay</i>	<i>in</i>	<i>váxt-i,</i>	<i>æz</i>	<i>bæ-ræsá-yme.</i>
until	tmw	this	time-OM	I	PU-arrive:PT-1S ₁
'By tomorrow this time, I will have arrived.'					

<i>áw=esan</i>	<i>obær</i>	<i>n-árd-e?</i>
water=3P ₂	to-out	NEG-bring:PT-PERF
'Haven't they extracted any water?'		

6.5.10. Past perfect and subjunctive perfect

All languages in the area have Past and Subjunctive Perfects, built on the formula: (punctual *bV-*) + Past Participle + Past/Subjunctive, resp., of AUX:be_j. Lāhijāni, however, replaces the Subjunctive Perfect with the Past Subjunctive. Most of Talyshi and nearby Tati varieties do not use *bV-* (§6.5.6.2) and elsewhere it is suppressed by preverbs or negatives. Mazanderani drops the final -*æ* of the participle in the Perfects.

Lāhijāni		Rashti		
	Past Perfect	Subjunct. Perf.	Past Perfect	Subjunctive Perfect
1SG	<i>bó-xord-ə bu-m</i>	(see Past Subjunctive, §6.5.11)	<i>bu-gúft-ə bu-m</i>	<i>bu-gúft-ə b-i-m</i>
2SG	<i>bó-xord-ə b-i</i>		<i>bu-gúft-ə b-i</i>	<i>bu-gúft-ə b-i</i>
3SG	<i>bó-xord-ə b-u</i>		<i>bu-gúft-ə bu-Ø</i>	<i>bu-gúft-ə be-Ø</i>
	'eat'		'say'	

Sāravi:

	Past Perfect	Preverbs	Subjunctive Perfect	Preverbs
1SG	<i>bæ-merd bi-me</i>	<i>dæ-či bi-me</i>	<i>bæ-mérd bu-(e)m</i>	<i>dæ-či bu-(e)m</i>
2SG	<i>bæ-merd bi (<i-i)</i>	<i>dæ-či bi</i>	<i>bæ-mérd bu-i</i>	<i>dæ-či bu-i</i>
3SG	<i>bæ-merd bi-ye</i>	<i>dæ-či bi-ye</i>	<i>bæ-mérd bu-e</i>	<i>dæ-či bu-e</i>
	'die'	'pick'	'die'	'pick'

Leriki:	Past Perfect		Subjunctive Perfect	
	Intransitive	Transitive	Intransitive	Transitive
1SG	<i>vit-áe b-im</i>	<i>hard-áe=m be</i>	<i>vit-áe bu-m</i>	<i>hard-áe=m bu</i>
2SG	<i>vit-áe b-iž</i>	used only fronted	<i>vit-áe bu-ž</i>	used only fronted
3SG	<i>vit-áe be-Ø</i>	<i>hard-áe=ž be</i>	<i>vit-áe bu-Ø</i>	<i>hard-áe=ž bu</i>
1NEG	<i>vit-áe nü-b-im</i>	<i>hard-áe=m nü-be</i>	<i>vit-áe nü-bu-m</i>	<i>hard-áe=m nü-bu</i>
	'run'	'eat'	'run'	'eat'

Māsulei:	Past Perfect		Subjunctive Perfect	
	Intransitive	Transitive	Intransitive	Transitive
1SG	<i>siáe-m=a</i>	<i>hérd=əm=a</i>	<i>siáe bù-bu-m</i>	<i>hærdáe=m bù-bu</i>
2SG	<i>sié-š=a</i>	<i>hérd=i=a</i>	<i>siáe bə-b-e</i>	<i>hærdáe=y bù-bu</i>
3SG	<i>si=a</i>	<i>hérd=əš=a</i>	<i>siáe bù-bu-Ø</i>	<i>hærdáe=š bù-bu</i>
1NEG	<i>siáe nim=a</i>	<i>hærdáe nim=a</i>	<i>siáe mè-bu-m*</i>	<i>hærdáe=m mè-bu*</i>
	'burn'	'eat'	*(for the negator morpheme, see 6.5.2, point 3)	

As mentioned above, all Perfect paradigms in Vafsi take the punctual marker *bV-* in addition to placing the word stress on the past root of the verb.

Vafsi: Intransitive

	Past Perfect	Subjunct. Perfect
1SG	<i>bæ-mærd-æ v-im</i>	<i>bæ-mærd-æ vue-ym</i>
2SG	<i>bæ-mærd-æ v-i</i>	<i>bæ-mærd-æ vue-y</i>
3SG	<i>bæ-mærd-æ v-e</i>	<i>bæ-mærd-æ vuæ-Ø</i>
1NEG	<i>næ-mærd-æ v-im</i> ‘die’	<i>næ-mærd-æ vue-ym</i>

Transitive

Past Perfect	Subjunct. Perfect
<i>b-im-rútt-æ v-e</i>	<i>b-im-rútt-æ vuæ-Ø</i>
<i>b-i-rútt-æ v-e</i>	<i>b-i-rútt-æ vuæ-Ø</i>
<i>b-is-rútt-æ v-e</i>	<i>b-is-rútt-æ vuæ-Ø</i>
<i>n-ís-rutt-æ v-e</i> ‘sell’	<i>n-ís-rutt-æ vuæ-Ø</i>

Examples of the Past Perfect

(110) Lāhijāni (LTXT)

<i>mualləm</i>	<i>n-áma</i>	<i>bu-Ø,</i>	<i>bÚ-šo-im</i>	<i>futbol</i>	<i>bazí</i>
teacher	NEG-come:PT	AUX:PT-3S	PU-go:PT-1P	soccer	play
<i>bù-kon-im.</i>				PU-do-1P	

‘The teacher had not come (so) we went (out) to play football.’

(111) Sāravi (SY: 66)

<i>čæn=ta</i>	<i>dez</i>	<i>ke</i>	<i>šeter-a=re</i>	<i>bæ-dezi</i>	<i>bi-ne,</i>
a.few=UNC	thief	SUB	camel-PL=RA	PU-steal:PT	AUX:PT-3P
<i>b-urd</i>	<i>bi-ne</i>	<i>sere</i>	<i>dele</i>		

PU-gone AUX:PT-3P house in

‘A few thieves who had stolen the camels had gone into the house.’

The Subjunctive Perfect is triggered when a subjunctivizing factor (a modal, a conditional clause, ‘perhaps’, ‘I hope’, etc.) acts on a Present Perfect: ‘it has ended’ > ‘it must have ended’, ‘if it has ended’, ‘perhaps it has ended’, ‘I hope it has ended’, etc.

(112) Sāravi

<i>[yε-n-e]</i>	<i>ve=re</i>	<i>ba-ut</i>	<i>bu-e.</i>
must-PR-3S	he-RA	PU-say:PT	AUX:SBJ-3S
‘He[must] have told him.’			

sayed *xord bu-em.*
 perhaps eat:PT AUX:SBJ-1S
 ‘Perhaps I have eaten (it).’ (SH:123)

(113) Leriki

gasti *čay zo-on omé bu-n.*
 perhaps his son-PL come:PT AUX:SBJ-3P
 ‘Perhaps his sons have come.’

Vafsi

sayæd *b-is-vattæ vuæ-Ø.*
 perhaps PU-3S₂-say:PT AUX:SBJ-3S₁
 ‘Maybe:s/he as said (so).’

(114) Lāhijāni (Perfect Subjunctive merges with the Past Subjunctive; NJ: 136)

bə *gəmon=əm* *ta amə bá-rəs-im,, dəro təmən=a*
to guess=1s₂ until we PU-arrive-1P harvest ended=CS
bə *b-i-Ø.*
 become:PT AUX:PT-IRR-3S
 ‘I guess that by the time we arrive, the harvest will have ended.’

(115) *agə* *nəhar=a bù-xord-ə b-in,, mu gursənə mon-əm.*
 if lunch=RA PU-eat:PT-PPL AUX:PT-3P I hungry stay-1s
 ‘If they have eaten the lunch, I will remain hungry.’ (NJ: 138)

6.5.11. Past subjunctives, conditionals, and other counterfactuals

Certain Caspian and Tatic varieties have special paradigms for counterfactual uses. Two basic Past Subjunctive types are formed by: 1) the Subjunctive marker *bV-* + the Imperfect built on the present stem; 2) the Past Participle + irrealis AUX in an analytic Perfect-like form. Most languages given here have more than one of these paradigms and they often overlap in usage with the Imperfect (conditional senses) or Past Perfect (irrealis senses). In some languages the latter two paradigms are virtually the only possibility for expressing irrealis or conditional notions (see also exs. 141–144).

One feature shared by types 1 and 2 is the irrealis *-i- ~ -æy*, also the marker of the Imperfect in many languages. It is used in both synthetic (Leriki: *b-æ-š-i-š*, 2s) and analytic Perfect-like forms (Lāhijāni: *b-əma b-i-Ø*, 3s). In the Past Subjunctive the Asālemi Imperfect *-i-* changes to *-e-*, possibly an alternate form of *-i-: -æy- > -e- > -i-*.

In Lāhijāni, Langerudi, Vājārgāhi, Rāmsari, and possibly Tonekāboni, all person and number in the Past Subjunctive have become neutralized to a single 3SG form. This neutralization seems to be optional in some varieties as my own

elicitation materials on Lāhijāni show only personal forms, e. g., (116) vs. (117) or (135) vs. (134).

What unites these various paradigms is their use in some or all of the following:

1. With past necessity when the Present vs. Past of ‘must’ is only a *morphological* but not a *semantic* distinction, e. g., *bayæd/bayéstí*, used only tense-neutrally;
2. After the Desiderative particle *kaš ~ kaške/ki* ‘would that, if only, I wish’;
3. In the protasis of counterfactual conditional sentences: ‘if I knew/had known’;
4. In a restricted zone, past modals and main verbs with control features generally require a Past Subjunctive in the subordinate verb (exs. 127, 128, see also §6.6);
5. Miscellaneous: ‘it would have been better to’; ‘I almost = it was close that I would (do)’; ‘I would have preferred/liked to have gone’; ‘as if it were’; etc.

The range of irrealis situations is a rather complex topic that has only been meagerly addressed in the literature. The present analyses must be taken as preliminary and provisory until our knowledge on these issues is furthered. In the case of multiple paradigms, for example, the exact use of each needs to be fully investigated.

6.5.11.1. The paradigms

Past Subjunctive

	Leriki I	Asālemi I	Māsulei I	
1SG	<i>b-áe-žæn-i-m</i>	<i>b-áe-vin-e-m</i>	<i>bá-vu-(y)m</i>	
2SG	<i>b-áe-žæn-i-ž/š</i>	<i>b-áe-vin-e-š</i>	<i>bá-vu-(y)-š</i>	
3SG	<i>b-áe-žæn-i-Ø</i> ‘hit’	<i>b-áe-vin-e-Ø</i> ‘see’	<i>bá-vu-(y)-Ø</i> ‘say’	
	Māsulei II (intr., ‘come’)	Harzani (MH:77) (intr., ‘sleep’)	Keringāni (intr., ‘sleep’, <i>ex141</i>)	Leriki II (intrans. ‘sit’)
1SG	<i>omá bì-b-i-m</i>	<i>hetæ=m b-i-Ø</i>	<i>het-en=aey*</i>	<i>níštae b-æ-bæ-ym</i>
2SG	<i>omá bì-b-i-š</i>	<i>hetæ=r b-i-Ø</i>	<i>het-i=aey</i>	<i>níštae b-æ-bæ-yž</i>
3SG	<i>omá bì-b-i-Ø</i> (trans., ‘send’)	<i>hetæ=y b-i-Ø</i> (tr., ‘close’, <i>ex139</i>)	?	<i>níštae b-æ-bæ-y</i>
3SG	<i>venné=š bì-b-i</i>	<i>bæsda=y b-i-Ø</i>	<i>haerd=š=aey</i>	<i>votæ=ž b-æ-bæ-y</i>
			*(alt: <i>het=aey=n</i>)	

	Leriki III Miller: 'Preterit Conditional' (intrans. 'sit')	Leriki IV 'Potential Conditional' (intrans. 'sit')	Asālemi II (intrans., 'die')	Asālemi III (intrans., 'die')
1SG	<i>ništæ b-æ-i-m</i>	<i>ništæ b-é-b-i-m</i>	<i>mærdæ b-e-m</i>	<i>mærdæ b-æ-b-e-m</i>
2SG	<i>ništæ b-æ-i-ž</i>	<i>ništæ b-é-b-i-ž</i>	<i>mærdæ b-e-š</i>	<i>mærdæ b-æ-b-e-š</i>
3SG	<i>ništæ b-æ-i-Ø</i> (trans., 'say')	<i>ništæ b-é-b-i-Ø</i> (trans., 'say')	<i>mærdæ b-e-Ø</i> (trans., 'see')	<i>mærdæ b-æ-b-e-Ø</i> (trans., 'see')
3SG	<i>votæ=ž b-æ-i</i>	<i>votæ=ž b-é-b-i</i>	<i>vindæ=š b-e</i>	<i>vindæ=š b-æ-b-e</i>
	Lāhijāni	Langerudi/ Rāmsari	Tonekāboni	Sāravi
1SG	(same as 3s)	(same as 3s)	?	<i>bæ-rut b-i bu-em~bu-m</i>
2SG	(same as 3s)	(same as 3s)	?	<i>bæ-rut b-i bu-i</i>
3SG	<i>b-ut-i b-i-Ø</i> 'say'	<i>bo-gut-i b-i-Ø</i> 'eat'	<i>bé-git-e b-i-Ø</i> 'take'	<i>bæ-rut b-i bu-Ø</i> 'sell'
			(Amini:114,19)	

6.5.11.2. The various irrealis usages: conditional, past subjunctive and others

Now each of the five situations listed above that require the Past Subjunctives and other irrealis forms are examined individually here and in §6.5.12, Conditionals:

1) Past Necessity

While the modal 'must' has Present and Past forms, both are often devoid of any real sense of tense. Time is encoded by the Present Subjunctive (§6.6) or Past Subjunctive (or Imperfect, Conditional, etc.) in the subordinate verb. Some languages, e. g., Leriki, Māsulei, only have one tense-invariable form of the modal 'must'.

(116) Lāhijāni (NJ: 175)

parsal xasi qənat=a larubi b-ud-i b-i-Ø.
last.year must:PT qənat-RA dredging PU-do:PT-IRR AUX:PT-IRR-3S
'We should have dredged the *qanat* (underground canal) last year.'

One speaker of Lāhijāni offered five alternate forms in example (117). In the short time I worked with him, however, he never produced a Past Subjunctive form as introduced above. I only discovered this tense when going through Jahāngiri's (2003) sentences, and then working with speakers of Langerudi and Ramsari.

- (117) *bayésti zutter ...*
must sooner ...

VARIANT 1	VARIANT 2	VARIANT 3	VARIANT 4	VARIANT 5
<i>b-útt-ə=bu-m</i>	<i>Ø-gútt-ə=</i> <i>bu-m</i>	<i>Ø-gútt-am</i>	<i>b-útt-enæ=</i> <i>bu-m</i>	<i>Ø-gútt-</i> <i>enæ=bu-m.</i>
PU-say:	IMF-say:	IMF-say:PT-1S	PU-say:	IMF-say:
PT-PPL=	PT-PPL=		PT-COND=	PT-COND=
AUX:PT-1S	AUX:PT-1S		AUX:PT-1S	AUX:PT-1S
PAST PERF. I	PAST PERF. II	IMPERFECT	CONDITIONAL	CONDITIONAL DUR.

‘You should have said (something) sooner.’

- (118) Leriki

bæ-pe ištæ jif=o b-áe-hæ-y-m.
PR-must self pocket=from SBJ-DUR-eat-IMF-1S₁

‘I had to eat out of my own pocket (i. e., on my own expenses).’

tí bæ-pe vot-áe b-æ-b-i-Ø.
you PR-must say-PPL SBJ-DUR-AUX:PT-IRR-3S
‘You should have said (something).’

- (119) Māsulei

be omæ bi-b-i-Ø. be b-u-y-š.

must come:PPL SBJ-AUX:PT-IRR-3S must SBJ-come:IMF-IRR-3S
‘He was supposed to come/to have come.’ ‘You were supposed to come.’
(Past Subjunctive II) (Past Subjunctive I)

Some languages distinguish between the Present and Past of the modal ‘must’ in both form and function and always use the Present Subjunctive for the dependent verb:

- (120) Vafsi

Present

ær-go áe-ss-ame dar-an dáe-pič-am.
DUR-want PU-go-1P₁ tree-OP PVB-wrap-1P₁
‘We have to go wrap the rees.’

Past

ær-goa ke nešanæ=s báe-d-a.
DUR-want:PT SUB address=3S₂ PU-give-2P₁
‘You needed to give him an address.’

- (121) Razajerdi

m-o *gu-m.* *m-os* *gu-m.*
 DUR-want say-1s₁ DUR-want:PT say-1s₁
 'I am supposed to say.' 'I was supposed to say.'

- (122) Kordkheyli Mazanderani

yek men dune vé-n-e há-d-i.
 one man rice must-PR-3S PVB-give-2s
 'You have to give (us) one *man* (6 kilos) of rice.' (Borjian: M2,10e)
véss-e pela bæ-xer-en.
 must:PT-3S rice PU-eat-3P
 'They were supposed to eat pilaff.' (Borjian: M2,14d)

In Sāravi, if the tense of the modal is neutralized (123), the Imperfect is used rather than the Past Subjunctive (so far as is known) in the subordinate verb:

- (123)
- vé-n-e b-ur-e. véss-e Ø-ši-ye.*
-
- must-PR-3S PU-go-3S must:PT-3S IMF-go:DUR-3S
-
- 'He has to go.' (SH: 118) 'He had to go.' (SH: 118)

2) The Desiderative particle *kaš(ke ~ ki)* 'would that, if only, I wish'

With *kaški* the Past Subjunctive and the Past Perfect are both common:

- (124) Lāhijāni

kaški mu b-út-ə bu-m.
 if.only I PU-say:PT-PPL AUX:PT-1s₁
 'I wish I had said (something).'

[Verb: Past Perfect]

kaški bæ-mærd-ə b-i-Ø.
 if.only PU-die:PT-PPL AUX:PT-IRR-3S₁
 'I wish I had died.' (NJ: 173)

[Verb: Past Subjunctive]

- (125) Māsulei

kaški omæ bi-b-i-Ø.
 if.only come:PPL SBJ-AUX:PT-IRR-3S₁
 'I wish he had come.'

[Verb: Past Subjunctive II]

Rāmsari

kaš tə=ra bó-gut-i b-i-Ø.
 if.only you=RA SBJ-say:PT-PPL AUX:PT-IRR-3S
 'I wish I had told you'

[Verb: Past Subjunctive I]

- (126) Leriki

kaški inglisjæ zinæ=m b-áe-b-i-Ø.
 if.only English know:PT=1S₂ SBJ-DUR-AUX:PT-IRR-3S₁
 'I wish I knew English.'

3) Protasis of past counterfactual conditional sentences

For the protasis in past counterfactual sentences, see §6.5.12, Conditionals.

4) Use after Past Modals (areally restricted)

In some languages, the Past Subjunctive is automatically used after all past modals and past main verbs with control features over the dependent verb. This is a larger areal feature that encompasses all of Talyshi, Koluri, N. Tati, as well as Colloquial Armenian and dialects, Neo-Aramaic, and Georgian:

- (127) Māsulei

Present Modal	Past Modal		
<i>xa-m</i>	<i>bÚ-š-om.</i>	<i>xu-m</i> ~ <i>xu-y-m</i>	<i>bá-š-i-m.</i>
want-1S ₁	SBJ-go-1S ₁	want:IMF-1S ₁	want:IMF-IMF-1S ₁
'I want to go'		'I wanted to go'	SBJ-go-IMF-1S ₁

- (128) Asālemi

<i>b-æ-pist=im=æ</i>	<i>bá-(h)ær-om*</i>	
ALL-DUR-want:PT=AUX:1S ₁ =AUX	SBJ-eat-1S ₁	
'I want to eat.'	*	(> [bárom])
<i>æ-pi=m</i>	<i>bæ</i>	<i>bá-æ-(h)ær-e-m*</i>
DUR-want=1S ₂	AUX:PT	SBJ-DUR-eat-IRR-1S ₁
'I wanted to eat.'	*	(> [bárem])

5) Other miscellaneous situations:

Other miscellaneous situations that require a Past Subjunctive include the following (among others): 'it would have been better *had I gone* (but I did not)'; 'you almost (fell, etc.) = it was close that you would (fall)'; 'I would have preferred *to go/to have gone*'; 'I would have liked *to go/to have gone*'; '(just) as if (it were)'; etc.

- (129) Rāmsari

eynaen xod-e má=re hæme pæšø bæ-ze b-i-Ø.
 just.as self-EZ I-RA all mosquito PU-hit:PT AUX:PT-IRR-3S₁
 'It was just as if mosquitoes had bitten me all over.'

- (130) Asālemi

kæm bæ-mænd-æ gənəst-æ b-i-m ~ bá-gən-um.
 little ?-stay:PT-3S₁ fall-PPL AUX:PT-IRR-1S₁ SBJ-fall-1S₁
 'I almost fell (lit: little remained, i. e., it was close, that I might fall)'

6.5.12. Conditional paradigms

In the apodosis of irrealis conditional sentences, EG, Central Caspian, and Harzani have a formal Conditional, distinct from other paradigms. These languages in fact allow multiple alternates for both the Conditional and the Past Subjunctive. The regular imperfect is also allowed in the apodosis as in (131–132).

Conditional Paradigms: ‘say’, ‘eat’

	Lāhijāni	Lāhijāni	
	Conditional I	Conditional II	
1SG	<i>Ø-gút-ena=bu-m</i>	<i>b-út-ena=bu-m</i>	
2SG	<i>Ø-gút-ena=b-i</i>	<i>b-út-ena=b-i</i>	
3SG	<i>Ø-gút-ena=bu-Ø</i>	<i>b-út-ena=bu-Ø</i>	
	Langerudi	Harzani (MH:79)	
	Conditional I, (II)	Condit. (ex. 139)	
	<i>gút(-en)-enæ=bu-m</i>	<i>bæ-b-e ber-in</i>	
	<i>gút(-en)-enæ=b-i</i>	<i>bæ-b-e ber-i</i>	
	<i>gút(-en)-enæ=bu-Ø</i>	<i>bæ-b-e ber-æ</i>	
	Rāmsari	Tonekāboni*	Tonekāboni
1s	<i>xord-enæ ba-m</i>	Conditional I	Conditional II
	<i>xár-e ba-m</i>	<i>xár-en-e ba-m</i>	<i>bó-xard-ene ba-m</i>
	Tonekāboni	Tonekāboni	Tonekāboni
	Conditional IV	Conditional V	Conditional III
1s	<i>xárd-e ba-m</i>	<i>xárd-en-e ba-m</i>	(cf. Past Perfect: <i>bó-xard-e ba-m</i>)

*Note that Tonekāboni I, II are built on the present stem, the others on the past stem. All Tonekāboni forms from Amini: 80-84.

(131) Tonekāboni (Protasis: Past Subjunctive, Apodoses: Imperfect; Amini: 114)

áge vi sér=e ab b-om-é b-i-Ø,
if his on=RA water PU-COME:PT-PPL AUX:PT-IRR-3S₁
jó=e tekan Ø-dæ-Ø.
barley=RA moving IMF-give:PT-3S

‘If the water had come up over it, it would have moved the rice grains around.’

(132) Sāravi (Protasis: Past Subjunctive, Apodoses: Imperfect; SH: 152)

aeger bæ-xord b-i bu-m, te=re Ø-got-eme.
if PU-eat:PT AUX:PT-IRR AUX:PT-1S you=RA IMF-say:PT-1S
‘If I had eaten (it), I would have told you.’

Since various languages allow some variation in both the protasis and the apodosis, it is not uncommon to find conditional sentences with the same tense in both clauses.

- (133) Lāhījāni (Protasis: Conditional I/Apodosis: Conditional I)
- aegær mašin rošen bo-na bu-Ø,, amo*
 if car lit become:PT-COND AUX:PT-3s we
šo-na b-im.
 go:PT-COND AUX:PT-1P
 'If the car would start, we would go.'

- (134) *agə Ø-donəst-əm,, t-æ Ø-gutt-am.*
 if IMF-know:PT-1S you-RA IMF-say:PT-1S
 'If I knew, I would tell you'
 Pro: Imperfect/Apo: Imperfect
aga Ø-gutt-i,, či bo-na bu-Ø?
 if IMF-said-2s what be:PT-COND AUX-3S
 'If you told me, what would happen?'
 Pro: Imperf/Apo: Cond I

- (135) *agə duz=a ná-šənaxt-ə b-i-Ø, či Ø-gud-i?*
 if thief=RA NEG-know:PT-PPL AUX:PT-IRR-3S what IMF-do:PT-2S
 'If you hadn't recognized the thief, what would you do/would you have done?' (NJ: 138) Protasis: Past Subjunctive/Apo: Imperfect

- (136) *agə bá-xast-ə bu-m,, xórd-ena bu-m ~*
 if PU-want:PT-PPL AUX:PT-1S eat:PT-COND AUX:PT-1S
 'If I had wanted to, I would have eaten (it).' (Pro: Past Perf./Apo: Condit. I), alternatively:

bó-xord-ə bu-m ~ Ø-xórd-ə bu-m ~ bó-xord-ena bu-m.
 PU-eat:PT-PPL AUX:PT-1S IMF-eat:PT-PPL AUX:PT-1S PU-eat:PT-COND AUX:PT-1S
 Past Perfect I ~ Past Perfect II ~ Conditional II (in order of preference)

- (137) Harzani (MH: 76)
- aegær bəsda=y b-i-Ø,, čök bæ-b-e ber-æ.*
 if close:PT=3S₂ AUX:PT-IRR-3S₁ good ALL-be-INF AUX:PT-3S₁
 'If he had closed (it), it would have been better.'

In those languages that have more than one possible verb form for one or both clauses of conditional sentences, e. g., Sāravi (132) and (138), it is most likely the case that there is a semantic difference, cf. 'If I knew (now), I would tell you' vs. 'If I had known (then), I would have told you'. Not enough information is available about these issues in any one of these languages and most languages either

have no available information or no analyses have ever been made for them. See, however, Paul (2011: 149–151) for a discussion of Past Counterfactuals in Talyshi dialects.

In some languages the Imperfect predominates in the apodosis of irrealis sentences:

(138) Sāravi

- aeger Ø-ši**, *ve=re Ø-di**.
 if IMF-go:IMF:2s he=RA IMF-see:PT:2s
 ‘If you went, you would see him.’ (SH: 152) *(*< Ø-ši-i, Ø-di-i, 2sg.*)

Rashti

- ager danést-i-m,, tð=ra gúft-i-m.*
 if know:PT-IMF-1s you=RA say:PT-IMF-1s
 ‘If I knew, I would tell you.’

(139) Leriki

- gilæ-y gám-æ vřræ b-æ-bæ-y,, œmæ=ní ištæ
 UNC-INDF warm-LINK place SBJ-DUR-be:PT-IRR we=also self
 vayœ œ-tosp-ín-i-mon.*
 there DUR-heat-CAUS-IRR-1P₁
 ‘If there were a warm place, we could (lit: would) warm ourselves there.’

(140) Keringāni (YSFN)

- hard=i=æy,, be-mard-e b-iš=e.
 eat:PT=2s₂=IRR ALL-die-INF AUX:PT-2s₁=AUX*
 ‘If you had eaten (it), you would have died.’

- aega het-en=æy,, mard-e b-in=e.
 if sleep-1S₂IRR die-INF AUX:PT-1S₁=AUX*
 ‘If I had slept, I would have died.’

Languages with no formal Past Subjunctive or Conditional generally use the Past Perfect or Imperfect in the various situations presented in this and the previous section:

(141) Vafsi

- xéyli dél-i=m œer-gòa œ-č-im* šæ:r
 very heart-OB=1s₂ DUR-want:PT DUR-go:PT-1s₁ city*
 ‘I would rather have gone to town.’ *(*< œt-si-im*)

- kaš zutær i-r-vattæ.
 if.only sooner 2s₂-DUR-say:PT*
 ‘I wish you had said sooner’

- (142) *áégæ dásd=i n-æj-jæ*, n-é-r-vušuz-i-a.*
 if hand=2s₂ NEG-DUR-hit:PT NEG-DUR-collapse-PASS-PT
**(n-æj-jæ < n-æd-zæ)*
 'If you hadn't touched (it), it wouldn't have collapsed.'
- (143) Alviri (HA: 172)
ægær et-omo-Ø, tæ=š e-di.
 if DUR-come:PT-3s₁ you=3s₂ DUR-see:PT
 'If he came, he would see you.'
- kaš et-omo-Ø.*
 if.only DUR-come:PT-3s₁
 'I wish he had come.'
- (144) Keläsi
m-óst-e zudtær me-še-m.
 DUR-want:PT-3s₁ sooner DUR-go:PT-1s₁
 'I should have gone sooner.'
- kaški zudtær a-m-vat.*
 if.only sooner PVB-DUR-say:PT
 'I wish I had spoken up sooner.'

6.5.12.1. Other conditional uses

Note the Past Subjunctive and the Conditional in the following Lāhijāni example:

- (145) *hær kæsə amra itow b-ud-i b-i, ti amra dušmən bo-na bo=Ø.*
 PAST SUBJUNCTIVE CONDITIONAL
 'Whoever you acted this way with would become an enemy with you.'
 (NJ: 123)

6.6. Modals, main verbs with control features

1) Modals ('want, can, must') and main verbs with control features (\pm coreferential subject) usually take a Subjunctive in the subordinate verb. Leriki also optionally allows an Infinitive (see ex. 160), the western Caspian, S. Talyshi and neighboring Tati less so, in Vafsi occasionally, and in others only a Subjunctive is used. Past Modals in Talyshi and N. Tati generally require a Past Subjunctive (§6.5.11.2).

- (146) Lāhijāni
ušon xá-n-ən æme=məra b-a-n.
 they want-PR-3P we:POSS=with PU-come-3P
 'They want to come with us.'

- xass-am bú-šu-m.
want:PT-1s PU-go-1s
'I wanted to go.'

(147) Sāravi
xa-n-e telafí há-kan-e.
want-PR-3s retaliation PVB-do-3s
'He wants to get even.' (SY: 62)

Ø-xas-e b-úr-e sæfer.
IMF-want:PT-3s PU-go-3s trip
'He wanted to go on a trip.' (SY: 64)

(148) šema beter bæled=eni báé-res-in.
you:PL better capable-2P PU-spin(thread)-2P
'You know how to spin better.' (SY: 64)

vé-n-e emsal aerusi há-kan-e.
must-PR-3s this.year marriage PVB-do-3s
'He must marry this year.' (SY: 58)

(149) Leriki
Subjunctive
pi-dæ=m ni tū-ni bi-voyand-im.
want-PR=1s₂ AUX:NEG you-ACC SBJ-send-1s₁
'I don't want to send you.'

Optative (§6.5.3)
az bæ-p-e mæktub bi-niivišt-o-m.
I FUT-want-3s₁ letter SBJ-write-OPT-1s₁
'I have to write a letter.'

(150) Vafsi
is-aer-go kay so-Ø?
3S₂-DUR-want where go:SBJ-3s₁
'Where does he want to go?'

tæmen aer-goa henra=s báé-sso-m
I:OB DUR-want:PT with=3s₂ PU-go-1s₁
'I wanted to go with him.'

(151) Khoini (YX: 179)
pist-aer=im báé-šu-m. pist-ae bi-m SBJ-šu-m.
want:PT-LOC=1s₁ SBJ-go-1s₁ want:PT-LOC AUX:PT-1s₁ PU-go-1s₁
'I want to go.' 'I wanted to go.'

- (152) Alviri (YA: 182)

em-e-gu-æ *bá-š-emæ.*
 1S₂-DUR-want-AUX:3S₁ PU-go-1S₁
 'I want to go.'

em-e-gust-æ *bá-š-emæ.*
 1S₂-DUR-want:PT-AUX:3S₁ PU-go-1S₁
 'I wanted to go.'

EG has a special negative form of 'can', lacking in WG: (Pres) *mánn(y)-* (Past) *mánnyyess-*, synonymous with the usual verb of Potentiality:

- (153) Lāhijāni

mu nó-ton-əm b-a-m. *mány-em* *bá-bin-əm.*
 I NEG-can-1s PU-come-1s can:NEG-1s PU-cut-1s
 'I can't come.' 'I can't cut (it).'

Vafsi modals 'can' and 'must' take the usual TAMs, may be negated, but do not take PAMs. (See §6.7 for a comparison of 'want' and 'must' formed from the same root.)

- (154)
- æz hæni n-æ-ču**
- v-á-yme.*

I else NEG-DUR-can PU-come-1S₁
 'I can't come (here) any more.' *(*< ná-æt-su*, root: *su* 'be able, can')

The use of the infinitive in the dependent verb is slightly more common with 'can' than with other modals or main control verbs. It may precede or follow the modal.

- (155) Lāhijāni

mu asla orə šo-n mánne-m
 I at.all there go-INF can:NEG-1s
 'I can't go there at all.'

Kelāsi
me-tan-ám bá-š-im(æ) ~ ši-æn.
 DUR-can-1S₁ PU-go-1S₁ go:PT-INF
 'I can go.'

- (156) Leriki

viškiy-e-dæ=š o-vašt-e.
 can-INF-LOC=2S₁ PVB-jump-INF
 'You can jump.'

æv hič næm-e zin-dæ ni-Ø.
 he none bend-INF know-LOC AUX:NEG-3S₁
 'He can't stoop over at all.'

6.7. Experiencer verbs

Experiencers (EXP) are marked by an Oblique case in Tatic and are coindexed by Set₂. The verbal element is an impersonal 3SG form of either a compound with the light verbs ‘be’, ‘come’, ‘become’, etc. or a simplex verb (restricted to one verb, ‘want/ love’). Generally, ‘want’ and ‘must’ are formed on the same verb root. In the sense of ‘must’ no clitics are used and the subject is in the Direct case.

(157) Vafsi

Experiencer

<i>tæmen</i>	<i>im-aer-gó</i>	<i>bæé-ss-om</i>
<i>esdæ</i>	<i>i-r-gó</i>	<i>bæé-sse-y</i>
<i>taní</i>	<i>is-aer-gó</i>	<i>bæé-sso-(æ)</i>
PRON:OB	PAM ₂ -DUR-want	PU-GO-PAM ₁
		‘I, you, s/he wants to go’

Non-experiencer

<i>æz</i>	<i>aer-gó</i>	<i>bæé-ss-om</i>
<i>tæ</i>	<i>aer-gó</i>	<i>bæé-sse-y</i>
<i>an</i>	<i>aer-gó</i>	<i>bæé-sso-(æ)</i>
PRON:DIR	DUR-want	PU-GO-PAM ₁
		‘I, you, s/he must go’

(158) Kafteji

<i>mən</i>	<i>me-gó=m</i>	<i>bé-š-im</i>
<i>tæ</i>	<i>me-gó=y</i>	<i>bé-š-iš</i>
<i>ja</i>	<i>me-gó=š</i>	<i>bé-š-i</i> (m.)
<i>ja</i>	<i>me-gó=š</i>	<i>bé-š-eæ</i> (f.)
PRON:OB	DUR-want-SET ₂	PU-GO-PAM ₁
		‘I, you, he, she wants to go’

<i>æz</i>	<i>me-gó</i>	<i>bé-š-im</i>
<i>tæ</i>	<i>me-gó</i>	<i>bé-š-iš</i>
<i>a</i>	<i>me-gó</i>	<i>bé-š-i</i> (m.)
<i>a</i>	<i>me-gó</i>	<i>bé-š-eæ</i> (f.)
PRON:DIR	DUR-want	PU-GO-SET ₁
		‘I, you, he, she must go’

Leriki (four types)

	Set ₂	Genitive	Genitive+Set ₂	Oblique
‘be cold’	<i>sárd=im=e</i>	<i>čimí sárd=e</i>	<i>čimí sárd=im=e</i>	<i>mí sárd=e</i>
(1/3SG)	<i>sárd=iž=e</i>	<i>čæy sárd=e</i>	<i>čæy sárd=iž=e</i>	<i>æy sárd=e</i>
‘be hungry’	<i>væší=m=e</i>	<i>čimí væší=e</i>	<i>čimí væší=m=e</i>	<i>mí væší=e</i>
(1/3SG)	<i>væší=ž=e</i>	<i>čæy væší=e</i>	<i>čæy væší=ž=e</i>	<i>æy væší=e</i>

- (159) Asālemi Māsulei
ari=m ni-æ. *xdeš=ər=æ* *b-am* *višir=i=æ?*
 late=1s₂ COP:NEG-3s₁ nice=2s₂=COP PR-COME:PT hungry=2s₂=COP
 'I'm not late.' 'you like' 'Are you hungry?'

Type one is well-known from Tehrani Persian (*særð=aem=e*). The fact that Set_2 in Talyshi encodes Ergative Agents and EXPs but not possessors is support for viewing the Set_2 PAMs encoding EXPs as Indirect Affectees rather than possessives.

With ‘want/love’, Set₂ is generally fronted to the Source noun or other element in the predicate. In Leriki, the dependent verb may be an infinitive:

- (160) Leriki
Fronted Set,
di *ti-[z]=æn* *om-e* *[pi-dæ]?*
with you=3s₂=also come:INF want-PR
‘Does he want to come with you, too?’

Unfronted Set ₂	
<i>piəso</i>	<u><i>pi-dæ-ž-e</i></u>
P.N.	want-PR=3s ₂
‘Does he want to go to Pirasora?’	š-e?

- (161) Kafteji
Fronted Set₂ Unfronted Set₂
 $\check{c}i\check{c}i=\check{s}on \quad me-go?$ $\check{c}i\check{c}i \quad me-go=\check{s}on?$
what=3P₂ DUR-want what DUR-want=3P₂

Oblique case, no Set₂
 $jon \quad \check{c}i\check{c}i \quad me-go?$
they:OP what DUR-want
‘What do they want?’

The Experiencer structure of the verb ‘love’ mirrors the Ergative PAM set-up. It should be noted that this alignment is found in all tenses in the Experiencer structure, but the Ergative is found only in the Past system:

The Set₂ encoding the Experiencer in the Light Verb type is obligatorily fronted to the NVE. Fronting farther to the left is not licensed:

- (163) Vafsi

<i>rīš-isbi-an</i>	<i>[dīv=esan]</i>	<i>[nǣ-tt-a-Ø]</i>	<i>æzin</i>	<i>aqáte</i>
beard-white-PL	face=3P ₂	NEG-DUR-come-3S ₁	thus	word
<i>bǣ-kær-ende.</i>				

DUR-do-3P₁
'The elders don't dare talk like this.'

Caspian languages generally do not coreference an overt EXP in the verb complex since the Set₂ PAMs we find in Tatic are not native to these languages. The EXP is instead marked with the (Direct/Indirect) Objective marker =ra.

- (164) Lāhijāni

<i>sada=jī</i>	<i>m-a</i>	<i>xoš</i>	<i>n-á-n-ə.</i>	<i>m-a</i>	<i>xow</i>	<i>a-n-e.</i>
plain=from	I-RA	nice	NEG-come-PR-3S	I-RA	sleep	come-PR-3S

'I don't like the plain one.' (NJ: 116) 'I am sleepy.'

- (165) Sāravi

<i>xæle</i>	<i>me=re</i>	<i>xoš</i>	<i>e-n-e.</i>	<i>šema=re</i>	<i>gærm=e?</i>
much	I=RA	nice	come-PR-3S	you:PL=RA	hot=COP:3S

'I like it very much.' (SH: 150) 'Are you hot?' (SH: 156)

On the other hand, depending on the given dialect or even speaker, these Set₂ PAMs "not licensed" for Caspian languages are clearly being borrowed from Persian, but so far only in peripheral domains, such as Experiencers and Reflexives (§5.9):

- (166) Lāhijāni

<i>amu</i>	<i>xow=emun</i>	<i>a-n-e.</i>	but: <i>in=a</i>	<i>xow</i>	<i>Ø-ama-Ø.</i>
we	sleep=1P ₂	come-PR-3S	he=RA	sleep	IMF-come:PT-3S

'We are sleepy.' 'He was (Imperfect) sleepy' (NJ: 176)

6.8. The verbs 'BE'

In a restricted cluster of languages in the western Caspian area, we find six distinct Be-verbs that express the usual Copula functions (equation, etc.), existence, location-at/in, and are then further subcategorized according to the animacy of the Subject. Included are: EG, WG, Rāmsari, Tonekāboni, S. Talyshi, and Kafteji-Kelāsi.

[Be]₁ is the Copular enclitic (Set₁ PAMs) corresponding to Present PAM suffixes of main verbs (here: 'to sleep'), sometimes with a slight variation in vowels:

Lāhijāni		Bābolsari		Leriki	
Present	Copula	Neg.	Present	Copula	Subjunctive
1SG <i>xósən-əm</i>	=əm	<i>nēm</i>	<i>xəs-eme</i>	=eme	<i>bíxit-İM</i>
2SG <i>xósən-i</i>	=i	<i>ni</i>	<i>xəsen-i</i>	=i	<i>bíxit-iš</i>
3SG <i>xósən-ə</i>	=ə	<i>ne</i>	<i>xəsen-e</i>	=e	<i>bíxit-i</i>

	Māsulei		Vafsi		Keringāni		
	Present	Copula	Past	Present	Copula	Subjunctive	Copula
1SG	<i>xəs-ám</i>	=em-æ	=em-a	<i>ədōss-om</i>	=im(e)	<i>fes-ene</i>	=ene
2SG	<i>xəs-í</i>	=eš-æ	=eš-a	<i>ədōss-i</i>	=i	<i>fes-iše</i>	=iše
3SG	<i>xəs-á</i>	=æ	=a	<i>ədōss-e</i>	=e	<i>fes-e</i>	=e

Be₁ Usage: Be₁ is used for equation with predicate nominals and adjectives. It can also replace all six other Be-verbs. Other tense forms are not considered enclitics.

Be₂ is a full form equivalent of the Copula formed on an independent root. It is only used in the Present affirmative in Caspian and most of Tati while other tenses and the negative forms are replaced by the corresponding forms of Be₁. In Talyshi and some Tati dialects, however, Be₂ Present also serves as the basis for the Past paradigms.

	Lāhijāni	Sāravi	Vafsi	Pres	Past
1SG	<i>ís-əm</i>	<i>hæss-eme</i>	<i>h-im(e)</i>	<i>híst-im</i>	<i>híst-im=be</i>
2SG	<i>ís-i</i>	<i>hæss-i</i>	<i>h-i</i>	<i>híst-iš</i>	<i>híst-iš=be</i>
3SG	<i>ís-ə</i>	<i>hæss-e</i>	<i>h-e</i>	<i>híst-e</i>	<i>híst-e=be</i>

Māsulei

Pres	Past
éss-em-æ	éss-em=a
éss-eš-æ	éss-eš=a
éss-æ	éss=a

[Be₂] Usage: Be₂ has three main uses: 1) general existence; 2) replaces Present of Be₁ (but not in N. Talyshi) in situations where the enclitic cannot occur: when stressed for focus or as a single, independent word. It may also replace Be₁ in unstressed uses; 3) languages with no ‘have’-verb use Be₂ in predicative possession (§6.9).

Existence

- (170) Lāhijāni Gandomābi
yə gúrg-i ís-ə har šow a-n-ə. *xoda éss-e*
 one wolf-INDF be₂-3s each night come-PR-3S God be₂-3s
 ‘There is a wolf (that) comes (around) every night.’ ‘God is./God exists.’ (YSFN)
 (NJ: 127)
- (171) Sāravi Leriki
æ-tta mærdi bi-ye. *tæbbi dišmen-on hüst-in.*
 one-UNC man be₁:PT-3S natural enemy-PL be₂-3P₁
 ‘There was a man.’ (SY: 64) ‘There are natural enemies.’

Standing independently, stressed (in the B examples only)

- (172) Lāhijāni Vafsi
 (A: *tu mariz n-i.*) B: *čére, ís-əm.* (A: *vəwsi ke n-e.*) B: *číra, h-e.*
 you sick be₁:NEG-2S why be₂-1s Vafsi EMF be₁-3S₁ why* be₂-3s
 ‘A: You aren’t sick. B: Yes, I am.’ A: He’s not a Vafsi. B: Yes, he is
 *('why' is the affirmative response to a negative question or statement.)

Simply replacing Be₁, without any triggering by information structure:

- (173) Lāhijāni
ki gú-n-ə u divonə ís-ə?
 who say-PR-3S he crazy be₂-3s
 ‘Who says he’s crazy?’
- Vafsi
vəws æzin qæssæbé-i h-e.
 P.N. thus village-INDEF be₂-3S₁
 ‘Vafs is this kind of village.’

Be₂ is replaced by Be₁ in negatives and all other tenses lacking in Be₂ (also ex. 171):

Be₃ is formed with the past participle of ‘stand’ + Be₁ in various corresponding tenses. In Māsulei, ‘stand’ has merged with ‘stay’ > *mænn-*, and it thus only has five Be-verbs.

	Lāhijāni		
	Present	Negative	Past
1SG	íssam	náyssam	íssæ bum
2SG	íssay	náyssay	íssæ bi
3SG	íssa-Ø	náyssa-Ø	íssæ bu-Ø

Māsulei		
Present	Negative	Past
<i>mænnem=a</i>	<i>mænné nim=a</i>	<i>mænnem=a</i>
<i>mænneš=a</i>	<i>mænné niš=a</i>	<i>mænnes=a</i>
<i>mænnæ-Ø</i>	<i>mænné ni=a</i>	<i>mænn=a</i>

Be₃ Usage: Be₃ encodes location/existence-at/in but only for animates.

- (175) Lāhijāni
 Location
kelās-e *no-e* *daebiristan* *íssæ* *bu-m*
 class-EZ 9-EZ high.school be₃ AUX:PT-1S
 'I was in ninth grade of high school'

Existence
muælləm *né-ysæ* *bu-Ø.*
 teacher NEG-be₃ AUX:PT-3S
 'The teacher was not around.'

Be₄ is formed with an older adposition *dər ‘in, at’ + Be₁ in all tenses and negatives, with or without a synchronic locative postposition. Be₄ in its narrow sense is found in WG, EG, Central Caspian, Southern Talyshi, and some neighboring Tati languages. On the other hand, it is found in its wider sense quite extensively beyond this area, as discussed under §6.8.1, ‘Diminishing Number of *Be*-Verbs’.

Lāhijāni			Māsulei		
	Present	Negative	Past	Present	Negative
1SG	<i>därəm</i>	<i>dánnyəm</i>	<i>dábum</i>	<i>dərím=a</i>	<i>dəryé nim=a</i>
2SG	<i>dári</i>	<i>dánni</i>	<i>dábi</i>	<i>dərís=a</i>	<i>dəryé niš=a</i>
3SG	<i>dárə</i>	<i>dánnyə</i>	<i>dábu</i>	<i>dərí=a</i>	<i>dəryé ni=a</i>

[Be₄ Usage:] Be₄ functions as location/existence-at/in for non-humans; uses with humans are quite complex and are not included here due to space limitations:

(177) Lāhijāni

Location-in	Existence-in
<i>čai yuri mien dár=a.</i> tea teapot in be ₄ =3s 'The tea is in the teapot.'	<i>yuri mien čai dár=a?</i> teapot in tea be ₄ 3s 'Is there (any) tea in the teapot?'

(178) Māsulei

Location-in
<i>pul če jiv-é delæ derí=a.</i> money his pocket-OB in be ₄ 3s ₁ 'The money is in his pocket.'

Existence-in
<i>deg-é dælæ pøla derí=a?</i> pot-OB in rice be ₄ 3s ₁ 'Is there any rice in the pot?'

[Be₅] is based on the past participle of 'put' + Be₁ in various corresponding tenses. Be₅ only refers to inanimates, and thus only occurs in the 3rd person.

Lāhijāni			Māsulei		
	Present	Negative	Past	Present	Negative
3SG	<i>hánna-Ø</i>	<i>n-ðnnæ-Ø</i>	<i>hánnæ-bu-Ø</i>	<i>noð-Ø</i>	<i>noð niæ</i>

[Be₅ Usage:] Be₅ indicates existence and location of inanimates.

(179) Lāhijāni

<i>ja:be*</i>	<i>mien va-n-e</i>	<i>pul hánna bu-n.</i>
box:REZ	in must-PR-3S	money be ₅ AUX:PT-3P
'There must be money in the box.' *(<i>ja:bá-a</i>)		
<i>pul ko hánna-Ø?</i>		
money where be ₅ 3s		
'Where is the money?'		

- (180) Māsulei Tōnekāboni
pəla deg-é delæ no-á? *quri kuji nia-*Ø?
 rice pot-OB in be₅-AUX:3s teapot where be₅3s
 'Is there any rice in the pot?' 'Where is the teapot?' (Stilo FN)

Be₆ is the verb ‘stay’ and is rare as a Be-verb as it seems to occur only in the subjunctive. Since it only pertains to inanimates, it only occurs in the 3rd singular.

Lāhijāni:	Present	Negative
3SG	bá-mon-ə	??

Be₆ Usage: Be₆ indicates existence in reference to inanimates. It is an alternate to the subjunctive of Be₅. I was only able to elicit one example for Lāhijāni, but it was spontaneously offered without any prompting for the use of this verb vs. Be₅. Parallel examples of this verb and its very specific usage also exist for Rashti.

- (181) *tin-ə** *yek* *češmě-i* *íra* *bú-mon-ə.*
can-3s one spring-INDF here PU-stay-3s
‘There could be a spring here.’ *(*tin-ə* ~ *tan-ə*)

6.8.1. Diminishing number of be-verbs

As we move out of the core area of this isogloss, the number of Be-verbs decreases to three. Mazanderani and Rudbāri have $\text{Be}_{1,2,4}$ —they also extend southward at least as far as Gaz and Esfahan (see Stilo 2007c, 2007b, resp.)—and in these languages, Be_4 subsumes many functions of the other verbs. For example, Be_4 in all these languages, in addition to inanimate nouns, regularly includes animate and human nouns in the sense of Existence/Location Existence-in/Location-in, generally avoided in WG/EG:

- (182) Sāravi
 α -*tta* *de-e* *dele* *de=ta* *berar* *dæ-i-ne.*
 one-UNC village-REZ in two-UNC brother be₄-PT-3P
 'In a village there were two brothers.' (SY: 58)

hiške me=je pirtær dæ-ni-ye.
 no.one I=from older be₄-NEG-3S
 'There isn't anyone here older than me.' (SH: 150)

- (183) *Tutkāboni*
færda xanæ dær-i? bæle, dær-om. maryanæ dæ-ni-æ?
 tomorrow house be₄-2s yes be₄-1s egg be₄-NEG-3s
 'Are you home tomorrow? Yes, I am.' 'Aren't there any eggs?'

6.9. ‘To have’ and ‘to have not’

In the Caspian area we find two types of predicative possession separated by a clear isoglossic line. The Caspian family, Southern Talyshi, most Tati languages, Persian, etc. have a transitive ‘have’ verb. To the west, Central/Northern Talyshi, Northern Tati, as well as Kurdish, Zazaki, Hawrami, Azeri, Udi, Georgian, and Semitic have periphrastic predicative possessive constructions formed with existence verbs (or particles/verboids). This line even separates closely related varieties, e. g., Southern Talyshi from Central and Northern Talyshi or Southern and some Central Tati from Northern and other Central Tati.

Languages of the Transitive *Have*-verb Zone

The predicative possessive construction is based on the existence verb Be₂ (§6.8). Some languages, including Iranian Azeri, however, through a process of “Have-Drift” (Stassen 2005: 475), have adapted somewhat to the ‘have’-verb zone by incorporating the enclitics into the verb to form an independent paradigm for the periphrastic verb:

Leriki (Northern Talyshi)

Present Affirm.	Negative	Past Affirm.	Pres. Perfect
<i>híst=im=e</i>	<i>ní=m=e</i>	<i>híst=im=be</i>	<i>báe=m=e</i>
<i>híst=e < =i=e</i>	<i>ní=e</i>	<i>híst=e=be</i>	<i>báe:< báe=e=e</i>
<i>híst=iž=e</i>	<i>ní=ž=e</i>	<i>híst=iž=be</i>	<i>báe=ž=e</i>

Colloquial Azerbaijani

Present Aff.	Past Affirm.
<i>var-İM=di(r)</i>	<i>var-İM=idi</i>
<i>var-IN=di(r)</i>	<i>var-IN=idi</i>
<i>var-İ=di(r)</i>	<i>var-İ=idi</i>

Most Tatic languages mark the possessor argument with the Oblique case and/or cross-reference the possessor NP with floating Set, clitics in the predicate.

In Leriki, the possessor is commonly marked with an adposition, an Oblique case (nouns), or Genitive case (pronouns) and Set, is optional with an overt Possessor:

- (187) Leriki
Alternative formations

<i>kitob=im</i>	<i>hest(=e).</i>	<i>(bæ)</i>	<i>mii=ro</i>	<i>kitob</i>	<i>hest=b-e.</i>
book=1s ₂	exist(=COP:3s ₁)	to	I:OB=for	book	exist=COP:PT-3s ₁
'I have a book.'			'I had a book.'		

(188) *či* *mæktæb-í* *mašin=iš* *hišt=e.*

POSS	school-OB	car=3s ₂	exist=COP:3s ₁
'The school has a car.'			

čimii *šæš* *gilæ* *æyil=im* *hišt=e.*

I:POSS	six	UNC	child=1s ₂	exist=COP:3s ₁
'I have six children.'				

Keringāni and Northern Talyshi also allow elements within the predicate other than the possesum to host the Set₂ PAMs cross-referencing the possessor:

- (189) *Leriki*
 $d=\alpha\text{emén}[\bar{i}] \quad \check{c}\bar{r} \quad qulloy \quad [h\bar{ist}=e]?$
 with-we=2s₂ what service EXIST=3s₁
 'What job/service do you have with us?'

Keringāni
 $haesæn-e \quad pinju=\check{s} \quad pæs \quad [hesta] \quad [biæ].$
 P.N.-OB fifty=3s₂ sheep EXIST COP:PT
 'Hassan had 50 sheep.' (YSFN)

- (190) Karganrudi (Central Talyshi)
hærči *servæt* *hest=əš* *b-e* ...
 whatever wealth EXIST=3S₂ COP:PT-3S₁
 'Whatever wealth she had, ...' (Guizzo: 64)

- | | | | |
|--|-------------------|------|--------------|
| Harzani | <i>hest=aem=æ</i> | here | <i>zöræ.</i> |
| EXIST=1S ₂ =COP:3S ₁ | three | son | |
| 'I have three sons.' | (AAK: 148) | | |

While examples are not numerous, we can also assume that the fronting of Set₂ PAMs in the predicative possessive is not obligatory.

6.10. Valency: transitivizing and detransitivizing

6.10.1. Causatives: transitivization (synthetic and analytic)

Morphological causativization in these languages mostly functions as a transitivizer and occurs only sporadically with a few *transitive* roots. Space only allows for a very brief discussion of valency changing and a short list of causativized verbs.

The verbs in Table 6E loosely progress in volitionality from the most unaccusative intransitive type base to the one most common transitive verbs that can be causativized, ‘to eat’, which is generally only used when associated with hospitality, i. e., ‘they fed us well’, but not for feeding a baby, etc. (See also exs. 81, 139, 214, 237, 290.)

The most common Transitivizing strategy is an analytic one, consisting of light verb constructions, most often formed with ‘do, make’. This strategy is an equipollent process discussed under Passivization in the next section. In addition, causativization of transitives is most commonly encoded analytically in this area, often using the verb ‘to give’ in the area under investigation. This type will not be further addressed here.

In Table 6E, the intransitive base can be identified simply by dropping the causative suffix after the root. In cases where causativization requires a phonological change in the root, both roots are given in the format of (preverb)-intransitive > transitive.

Table 6E: Morphological causative formation

	Leriki	Asälemi	Mäsume	Vafsi	Lähijäni	Säravi
rot > let	<i>pix-ovn-</i>			<i>gend-en-</i>		
rot						
blaze, be lit	<i>væš- > voš-n-</i>		<i>væš- > ?</i>	<i>ves-en-</i>		
boil	<i>gil-ovn-</i>	<i>gəl-avən-</i>	<i>gəl-amən-</i>	<i>viš-en-</i>		
burn	<i>tæv > tov-n-</i>		<i>si > su-in-</i>	<i>suz-en-</i>		<i>suz-end-</i>
break	<i>si- > sī-ndin-</i>	<i>čæk->cok-ən-</i>	<i>čæk- > čak-ən-</i>	<i>æšk-en-</i>	<i>šk-en-</i>	<i>ešk-end</i>
tear, rip	<i>zīr-n-</i>		<i>ve-zer-amən-</i>	<i>æzerr-en-</i>		<i>boss-end-</i>
stick	<i>dæ-čik- > čok-n-</i>	<i>ar-čək- > čək-ovn-</i>	<i>čik-amīn-</i>	<i>dæ-duss-en-</i>	<i>du-čək-an-(*Rashti)</i>	<i>de-mass-end-</i>
rise, get up			<i>ez-amən</i>	<i>or-eyz-en-</i>	<i>vi-ris-on-</i>	<i>her-ess-end-</i>
sleep	<i>hit-ovn-</i>	<i>xəs-on-</i>	<i>xəs-əmən-</i>	<i>hoss-en-</i>	<i>xus-on-</i>	<i>xos-end-</i>

	Leriki	Asālemi	Māsulei	Vafsi	Lāhijāni	Sāravi
laugh	<i>sür-ovn-</i>	<i>xur-an'</i>	<i>xænd-amən-</i>	<i>xænd-en-</i>		
go around	<i>næv- > nov-n-</i>		<i>gærð-ən-</i>	<i>gærð-en-</i>	<i>gærð > gard-æn⁴</i>	<i>gerd > gard-en</i>
arrive > deliver	<i>ræs- > ros-n-</i>	<i>ræs- > ros-en-</i> ²	<i>a-res-amən-</i>	<i>ræs-en-</i>	<i>ræs-on-</i>	<i>res-end-</i>
fear	<i>tars->tors-İN-</i>	<i>tær-s-on-</i> ³	<i>tær-s-amən-</i>	<i>tær-s-en-</i>	<i>tær-s-ən-</i>	<i>tær-s-end-</i>
run	<i>vit-ovn-</i>		<i>tell-amən-</i>	<i>dæv-en-</i>		<i>tæj- > taj-en-</i>
jump	<i>vašt- > vošt-İN-</i>	<i>væz- > vaz-Ən-</i> ¹	<i>væz- > vaz- amən-</i>		<i>væz-en-</i>	
graze ⁵	<i>čard- > čord- on- ~ čo-vən-</i>	<i>čær- > čovən-</i> ²	<i>čær- > čar-ən- ~ čar-əmən-</i>		<i>čerr-en-</i>	
eat	<i>hard-ovn-</i>	<i>hær-an'</i>	<i>xor-amən-</i>	<i>hor-en-</i>		

1. Yarshater:1996; 2. Kargānrudi, Guizzo; 3. Tulārudi, Guizzo; 4. Vājārgāhi, Amirzādeh; 5. See ex. 214.

(191) Lāhijāni

Intransitive Base Verb

bu-šo-Ø ta bə-rəs-e arbáb-ə xonə.
 PU-go:PT-3S till PU-arrive-3S boss-REZ house
 ‘She went until she arrived at the boss’s house.’ (NJ: 143)

Transitivized Equivalent

tu i nam-á un=a rəs-ən-en-i?
 you this letter-RA he=RA arrive-CAUS-PR-2S
 ‘Will you deliver this letter to him?’ (NJ: 95)

(192) *bə-da amə rahat bú-xus-im. zak-on=a bú-xus-on.*

PU-give we comfortable PU-sleep-1P child-PL=RA PU-sleep-CAUS
 ‘Let us sleep peacefully.’ (NJ: 156) ‘Put the children to sleep.’
 (NJ: 149)

(193) Sāravi

taze hemdige=jem báe-resi-ne.
 new one.other-JE PU-arrive:PT-3P
 ‘They just ran into each other.’ (SH: 149)

me selam=re še berar=jem báe-res-en.
 my hello-RA self brother=JE PU-arrive-CAUS
 ‘Send my greetings to your brother.’ (SH: 156)

(194) Leriki

institut-í iminji kurs oræxæ-y.
 institute-OB first course end-PT-3S₁
 ‘The institute’s first course ended.’

johil-í medinstitut=íš orox-n-i-yæ.
 youth-OB med.school=3S₂ end-CAUS-PT-PPL
 ‘The young man finished med school.’

(195) Māsulei

<i>kərə Ø-si-ám.</i>	<i>kərə Ø-su-in-ám.</i>
PROG PR-burn-1s	PROG PR-burn-CAUS-1s
‘I am burning up (e. g., with fever).’	‘I am burning (something).’

(196) Vafsi

<i>gá-ye piš-ær-gerd-énde.</i>	<i>ga-an piš-gerd-en!</i>
COW-PL PVB-DUR-return-3P ₁	COW-OP PVB-return-CAUS
‘The cows are coming back.’	‘Take/bring the cows back.’

6.10.2. “Passives”: detransitivization (unaccusatives, inchoatives, etc.)

Valency-reducing strategies in simplex verb roots form two isoglosses in the area: 1) the synthetic type using various verbal suffixes, and 2) the analytic type formed with the past participle and the auxiliary ‘be(come)’. Most dialects of Tati and Southern and Central Talyshi have a synthetic formation. Northern Talyshi, Northern Tati, and all Caspian and Tatoid languages, however, only have analytic formations.

Synthetic passive markers, present/past

Hezārrudi	Kajali	Asālemi	Māsulei	Chāli	Alviri	Vafsi
-ax-/ax-est	-is-/ist(e)~ista, F	-i-/i-st-	-i-/i-áé	-i-/i-æs(s)-	-i-/i-æst	-uæ-/i-a

These voice paradigms incorporate a much wider range than the traditional label of “Passive” in these languages, such as unaccusative or inchoative, e. g., ‘the bud opened’, ‘the wall collapsed’, ‘the cup broke’, etc. As the examples below show, “Passive” is at times appropriate (‘be/get eaten/sold/thrown’), while in other cases no passive sense is implied at all (‘spills out’, ‘tears’). (See also exs. 15, 84, 142, 330.)

(197) Vafsi

- éger vəws há-wešaz-uæ-Ø, ... xun báé-riz-i-a-Ø.*
 if P.N. PVB-open-PASS-3S₁ blood PU-pour-PASS-PT-3S₁
 'If Vafs (village) opens up ...' 'The blood spilled out.'

(198) Hezārrudi (YT:465)

- hærči mi-kær-em, æm ni-mi-zær-ax-e.*
 whatever DUR-do-1S₁ this NEG-DUR-tear-PASS-3S₁
 'No matter what I do, this will not tear.'

ču a-škæj-ax-est-Ø.
 wood PVB-split-PASS-PT-3S₁
 'The wood split open.'

(199) Kajali (KJ: 283)

- divar a-škaj-is-t-Ø tænaf æd-ruj-is-t-Ø.*
 wall PVB-split-PASS-PT-3S₁ rope PVB-throw-PASS-PT-3S₁
 'The wall split open (and) a rope was thrown down.'

N. Talyshi, N. Tati and all of the Caspian family form the Passive analytically.

(200) Lāhijāni (NJ: 89)

- tumóm-ə γəza bú-xord-e bu-bo-Ø.*
 all-EZ food PU-eat:PT-PPL PU-AUX:PT-3S
 'All of the food was eaten.' (see also ex. 135)

Sāravi (SY: 40)

- parče-a bæ-rut-e bæ-i-ye.*
 fabric-PL PU-sell:PT AUX-PT-3S
 'The fabrics were sold.'

(201) Kalāsuri (KL: 279)

- em izem onjar-e ber-ondæ ni-Ø.*
 this firewood mince-PPL be:INF-LOC.PR? AUX:NEG-3S₁
 'This firewood won't get cut/will just not cut.'

6.10.3. Light verb constructions and change of state

Simplex verb roots are clearly in the minority in this area. My current Vafsi glossary consists of 138 simplex and 534 compound roots, but there are probably more simplex roots to be identified. The majority of transitive/causative vs. intransitive/passive pairs use equipollent strategies by which Non-verbal Elements (NVEs), esp. adjectives and nouns, are combined with Light Verbs, esp. 'make' vs. 'be(come)', resp.

Compare the following pairs derived from NVEs: (Leriki) *kok* 'fat' > *kok kard-*

(fat-make) ‘fatten’ vs. *kok b-* (fat-become) ‘get fat’; (Lāhijāni) *sefid* ‘white’ > *sefid gud-* (white-make) ‘whitewash (wall), bleach (clothes)’ vs. *sefid bo-* (white-become) ‘blench/pale (of face), turn white (hair)’. Various lesser verb pairs of this type use other sets of Light Verbs, e. g., Sāravi: *šekes he-da-* (defeat-give) ‘defeat’ vs. *šekes xord-* (defeat-eat) ‘be defeated’; Vafsi: *ra veyn-* (road-throw) ‘send off’, *ra gen-* (road-fall) ‘start off (going, -TR)’; Leriki: *bæ æmæl vard-* (to-action-bring) ‘realize, implement, execute’, *bæ æmæl omæ-* (to-action-come) ‘materialize, be realized’.

[Change of State]: In an isogloss that includes Gilaki, Central Caspian, S. Talyshi and all areas as far south as Vafsi, those Causative/Passive pairs formed with the LVs ‘do’ and ‘become’ (but not other LVs) use the “Change of State” (CS) morpheme =*a*, encliticized to the NVE (usually an adjective, occasionally a noun). In punctual tenses either the punctual morpheme *bV-* or, less commonly, the =*a* may be deleted, but both may also occasionally occur together. In most cases, the CS or the punctual *bV-* is the only way to distinguish ‘be’ from ‘become’, (since ‘be’ does not take the *bV*: Lah. *bu* ‘he was’ vs. *bú-bu* ‘he became’), e. g., (205A) in Māsulei. Languages marked with an asterisk in this section are outside this isogloss and lack this morpheme.

Causatives and Passives of Non-Simplex Roots: Light Verb Constructions

- (202) Lāhijāni (NJ: 176)

Unaccusative, ‘Passive’ (Unacc.)

... *harf bə-zən-i,, bidár=a nə-bə-n-ə.*
speech PU-LV-2s awake=CS NEG-be-PR-3s
(If) you talk, he won’t wake up.’

Causative > Transitivized (Trans.)

ín-ə bərarzə ín=a bidár=a gud-ə.
he-REZ nephew he-RA awake=CS do:PT-3s
'His nephew woke him up.'

- (203) *Sāravi (SH: 153–154)

yæx u bæ-yy-e. še ja=re ævez ha-kord-ene.
ice water PU-become:PT-3s self place=RA change PVB-do:PT-3P
'The ice melted.' (Unacc.) 'They changed their place.' (Trans.)

- (204) *Leriki

vot-dæ=n vašag=æn pæydo b-æ-Ø
say-PR-3P₁ lynx=also visible be:PT-PERF-3S₁
'They say lynx were found (back then).' (Unacc.)

ro-y pæydo kard-e zin-dæ-ni-Ø.
road-OB visible do-INF know-PR-NEG-3S₁
'He can't find the road.' (Trans.)

(205) Māsulei

<i>ešτæ</i>	<i>pa</i>	<i>čel=a</i>	<i>bæ-Ø.</i>	<i>reč=a</i>	<i>kærd-əm=a.</i>
your	foot	muddy=cs	be:PT-3s ₁	straight=cs	do:PT=1s ₂ =AUX
'Your feet got muddy.' (Unacc.)					'I fixed it.' (Trans.)

(206) Vafsi

<i>kar</i>	<i>soay</i>	<i>támb=a</i>	<i>r-bù-Ø.</i>
work	tomorrow	finished=cs	DUR-become-3s ₁
'The work will be finished tomorrow.' (Unacc.)			
<i>kár=om</i>	<i>zu</i>	<i>támb=a</i>	<i>r-kær-om.</i>
work=1s ₂	early	finished-CS	DUR-do-1s ₁
'I'll finish my work early.' (Trans.)			

Note that the adjective plus 'be' is the base from which the equipollent Causative-Passive pairs are derived. Since 'be' has no Change of State sense, =a is not licensed.

(207) Vafsi

Stative	Unaccusative		
<i>čera</i>	<i>rušen</i>	<i>be.</i>	<i>čera</i> <i>rušen=a</i> <i>we.</i>
lamp	lit	was	lamp lit=cs became
'The lamp was on.'			'The lamp went on.'

Transitive

<i>čera</i>	<i>rušen=om=a</i>	<i>kærd.</i>
lamp	lit=1s ₂ =cs	did
'I turned the lamp on.'		

7. Alignment and marking of core arguments: nouns, pronouns, PAMs

7.1. Alignment systems and differential object marking (DOM)

Most languages of the area have DOM by which salient Ps—definite or specific nouns (including indefinite specific, see ex. 83; Vafsi also requires animacy)—are Accusative-marked. Non-salient Ps remain in the unmarked base form (SG, see §5.6).

In *RA*-marking languages—Caspian, Tatoid, Razajerdi, and Harzani—*Nominative-Accusative* alignment occurs in all tenses. Alviri has no case or P-marking. Tatic split alignment is discussed in §7.1.3. Harzani presents an unusual case: it has retained the formal Direct and Oblique cases of nouns (now reserved for possessive marking, §5.4.2) but has =RA-marking of Patients and Recipients in all tenses.

7.1.1. Nominative-accusative alignment, all tenses; =ra marking, DOM

- (208) Lähijāni (NJ: 156)

DOM: Salient, marked	Non-salient, unmarked
<i>mury=a</i> <i>xa-n-i?</i>	<i>mury</i> <i>xa-n-i?</i>
chicken=RA want-PR-2S	chicken want-PR-2S
‘Do you want the chicken?’	‘Do you want (a) chicken?’

- (209) Sāravi (SY: 62)

<i>ven-e</i> <i>mimun=re</i> <i>kúš-em-be.*</i>	<i>parsal</i> <i>ve</i> <i>mimun</i> <i>dašt-e.</i>
he-REZ guest=RA kill-PR-1S	last.year he guest have:PT-3S

‘I will kill his guest.’ *(*<kúš-en-me*) ‘Last year he had a guest.’

- (210) Rudbāri: Tutkāboni

<i>gusfænd=æ</i> <i>bə-frúxt-om.</i>	<i>gusfænd</i> <i>bə-frúxt-om.</i>
sheep=RA PU-sell:PT-1S	sheep PU-sell:PT-1S

‘I sold the sheep.’ ‘I sold a sheep.’

- (211) Tāleqāni: Gurāni

<i>der-im</i> <i>nan=e</i> <i>dé-m-bænd-im.</i>
BE ₄ .1P bread=RA PVB-DUR-bake-1P

‘We are baking the bread.’

<i>der-im</i> <i>nan</i> <i>dé-m-bænd-im.</i>
BE ₄ .1P bread PVB-DUR-bake-1P

‘We are baking bread.’

- (212) Harzani (see also ex. 280)

Tati:=RA-marking, Present

<i>te</i> <i>kuhurzani</i> <i>læv=e</i> <i>čök</i> <i>zunest-æ.</i>
you P.N. speech=RA good know-LOC?

‘You know Harzani well.’ (AAK: 108)

=RA-marking, Past

<i>mæn</i> <i>en</i> <i>ödæmi=re</i> <i>yær=mæ</i>
I this person=RA hit:PT=1S ₂

‘I hit_{past} this person.’ (MH: 308)

- (213) Razajerdi (RA-marking; YSFN)

<i>bez=æ</i> <i>bær</i> <i>ku.</i>	<i>verg-á=ræ</i> <i>köšt-em.</i>
goat=RA take mountain	wolf-PL=RA kill:PT-1S ₂

‘Take the goat to the mountain.’ ‘I killed the wolves.’

- (214) Alviri (no Accusative marking, Neutral alignment)

u hætmæn čarva e-čær-enden-æ.
he certainly sheep DUR-graze-CAUS-3S₁
'He will definitely graze the sheep.' (HA: 174)

aesb=eš deræxt=eš bæst.
horse=3S₂ tree=3S₂ tie:PT
'He tied his horse to a tree.' (HA: 310)

7.1.2. Case-marked objects in Tatic languages: present system, DOM

Most forms of Tatic have split alignment with Accusative (Oblique case P-Marking) in the Present. There are more complex alignments in the Past, treated in §7.1.3.

- (215) Leriki

DOM: Salient, marked	Non-salient, unmarked
<i>ov-i tát kárd-æ=š.</i>	<i>ov=aen bárd-æ=š</i>
water-OB warm do-PR=2S ₁	water=also take-PR=2S ₁
'You heat up the water.'	'You take some water, too.'

- (216) Māsulei

zoæ řsellay-é pi-er-ə. i-llæ řsellay mæ=ra b-oær-æ.
boy whip-OB PVB-take-3S₂ one-UNC whip I=for SBJ-bring-2P₁
'The boy picks up the whip.' 'Bring me a whip.' (LM: 46, 32)
(LM: 46, 34)

- (217) Vafsi (NB: Salient nouns in Vafsi must be both definite/specific and *animate*)

xoda tani laze-y ná Ø-dir-è.
God his son-OM keeping SBJ-have-3S₁
'May God preserve his/her son.'

xoda tæmen lazæ há-do-Ø.
God I:OB son PVB-give-3S₁
'May God give me a son.'

- (218) Gandomābi (YSFN)

æjem guspænd-é bé-bæ!
that:OB sheep-OM PU-take.away
'Take that sheep away!'

qessab guspænd sær a-ber-e.
butcher sheep head PVB-cut-3S₁
'The butcher slaughters a sheep.'

- (219) Dikin Marāqeи
das-aer=et mi-gir-en. man pikæ æsif á-gir-em.
 hand-OB=2s₂ DUR-take-3P I:OB want apple PVB-take-1s₁
 'I'll take your hand.' 'I want to get/buy an apple.'

7.1.3. Past tense argument alignments in Tatic

DOM is irrelevant in Ergative alignment since Ps, as Absolutives, are not marked. Vafsi (in two of its three Past patterns), Māsulei, and some other Tatic varieties have Double Oblique alignment and DOM thus applies to Past and Present Ps equally here.

Only nominal alignment is addressed in this section. Pronominal alignment is covered in §7.2. Set₂ PAMs coindexing Past Agents are generally optional in Tatic. Subjects (S) of intransitive verbs are disregarded in the schemata below. The cell for Set₁ (Direct) PAMs coindexing the P is marked [3] if it is restricted to the 3rd person. If the language does not distinguish gender, the genders are crossed out and should be considered common singular.

7.1.3.1. Ergative alignment of nouns and PAMs

The best examples of canonical Ergativity in Tatic are Dikin Marāqeи and Vafsi III. In the former this pattern is predominant, while it is only minor in Vafsi (see Vafsi I and II below). This pattern in both languages also retains coindexing of P in the verb in all persons. See §6.5.6.2 for the Objective conjugation in the past tense in Dikin Marāqeи.

	NP			PAM				NP			PAM				
	Agent			Patient			A P		Agent			Patient			A P
	m	f	p	m	f	p		m	f	p	m	f	p		
DIR				+	+	+					+	+	+	[3]	
OB	+	+	+				+					+		+	

- (220) Dikin Marāqei
killig-é *æ* *zolleg-ín=eš* *vínd-aen.*
 girl-OF that boy-DP=3S₂ see:PT-3P₁
 'The girl saw those boys.'

man=di *æ* *jænek* *vínd-ian.*
 I:OB=also that woman see:PT-3SF₁
 'I saw that woman, too.'

- (221) Kajali (KJ: 280)

<i>værg-on</i>	<i>pæs-e</i>	<i>be-hærd-end.</i>
wolf-OP	sheep-DP	PU-eat:PT-3P ₁
'The wolves ate the sheep _{pl.} '		
<i>mærdæk-i</i>	<i>žænég-i-æ</i>	<i>be-hind-a.</i>
man-OM	woman-INDEF-DF	PU-see:PT-3SF ₁
'A man saw a woman.'		

Khoynarudi (N. Tati) occasionally marks the Past Agent with the Direct case, but this pattern seems to be marginal and not consistent.

	NP			PAM			NP			PAM		
	Agent		Patient	A	P		Agent		Patient	A	P	
	m	f	p	m	f	p	m	f	p	m	p	
DIR	(+)	(+)		+	+		[3]					
OB	+	+	(+)	(+)	+			+		+	+	

- (222) Khoynarudi (YSFN)
 Past Agent in Direct case
 $em \quad ferqaeči-é \quad sér=čun \quad duræ=pe.$
 this P.N.-DP head=3_{P₂} give:PT=PPB
 ‘These *Fergachis* spoke up (lit: raised their heads).’

Past Agent in Oblique case
ferqæći-ín em=čün šahar estéræ.
 P.N.-OP this=3_{P₂} city take:PT
 'The *Fergachis* took this city.'

Northern Talyshi has Ergative alignment in nominal arguments but does not coindex Ps in the verb in any tense. Ergativity is neutralized in plural forms (223) since case is not distinguished in the N. Talyshi plural.

- (223) *ramez-i k'amil=iž-aen pe-gi š-e-Ø.*
 P.N.-OB P.N.=3S₂=also PVB-take go-PT-3S₁
 'Ramez took Kamil, too, and left.'

žen-on mol-on=ižon dušæ=y.
 woman-PL cattle-PL=3P₂ milk:PT=AUX
 'The women milked the cows.'

7.1.3.2. The double oblique alignment of salient NPs

Cháli has both Ergative and Double Oblique alignment with nominal arguments in the Past system, but the Ergative construction has no co-indexing of Ps in the verb:

- (224) Oblique-Direct (Ergative)

<i>œmir</i>	<i>aerselán-e</i>	<i>kæmær=eš</i>	<i>de-bæst.</i>
P.N.	P.N.-OM	belt=3s ₂	PVB-TIE:PT
'Amir Arselan fastened his belt.' (ST: 105, 16)			

Double Oblique (+ DOM)

værziær-o bærr-on=ešo b-ašind.

farmer-OP spade-OP=3P_s PU-throw:PT

‘The farmers threw down (their) spades.’ (ST: 76, 106)

Māsulei seems to have exclusively Double Oblique Alignment (with DOM):

- (225) *œ mərdæk-ón əštə p-ær venn=a.*
 that man-OP your father-OB see:PT=AUX:PT
 ‘Those men had seen your father.’

mən əm-é t=ræ voerd=əm=æ.
 I-OB this-OB you=for bring:PT=1S₂=AUX
 'I brought this for you.'

7.1.3.3. Vafsi past alignments

It should be noted that saliency in Vafsi requires both Definiteness/Specificity and Animacy. (See Dikin Marāqeи above for the schema of the Vafsi Ergative.)

Vafsi	NP			PAM		NP			PAM			
	Agent		Patient	A	P	Agent		Patient	A	P		
	m	f	p	m	f	p	m	f	p			
DIR												
OB	+	+	+	+	+	+	+	+	+	(+)		
	Double Oblique I (SOV)						Double Oblique II (OSV)					

The Past in Vafsi has three different Alignments for marking core arguments:

- 1) The most common alignment by far is the Double Oblique I. Various strategies help to disambiguate the two Oblique-marked arguments as discussed below.
- 2) Vafsi III (Ergative), requiring OSV order, has become rather marginalized.
- 3) Also rare is the Double Oblique II (Vafsi II), also obligatorily OSV.

Double oblique I (SOV)

As mentioned, this construction is by far the most frequent of the three Past system alignments in Vafsi. Due to DOM only salient Ps are Oblique-marked:

- | | |
|--|---|
| (226) Salient Ps | Non-salient Ps |
| <i>tæmen gulle-y=m báe-ruttæ.</i>
I:OB calf-OM=1s ₂ PU-sell:PT | <i>tæmen yey gullæ=m báe-ruttæ.</i>
I:OB one calf=1s ₂ PU-sell:PT |
| 'I sold the calf.' (animate, definite) | 'I sold a calf.' (animate, indefinite) |

The P in Double Obliques I/II is also optionally co-indexed in the verb with Set₁:

- | | |
|---|------------------|
| (227) P not co-indexed | Set ₁ |
| <i>belæxæræ tini kell-i=s hágirættæ.</i>
finally he:OB girl-OF ₂ =3s ₂ PVB-take:PT | Set ₁ |
| 'Finally he married the girl.' | Set ₁ |

P co-indexed

- | | |
|------------------|---|
| Set ₁ | P co-indexed |
| Set ₁ | <i>esðæ zarru-án=i ye-dí-e-nde?</i>
you:OB child-OP=2s ₂ PU-see:PT-PERF-3P ₁ |
| Set ₁ | 'Have you seen the children?' |

Double oblique II (OSV, no agreement with A or P)

- | | |
|---|------------------|
| (228) | Set ₁ |
| <i>hæzíri šo do dane kærg-án=oan luás-i yáe-værdæ.</i>
yesterday night two UNC chicken-OP=1P ₂ fox-OM PU-bring:PT | Set ₁ |
| 'Last night the fox carried off two of our chickens.' | Set ₁ |

Vafsi III: The Ergative Construction (OSV, agreement with P)

- (229) *in dár-e æhmæd-i par=es dðé-nnia-nde*
 this tree-DP P.N.-OM last.year=3S₂ PVB-put:PT-3P₁
 ‘Ahmad planted these trees last year.’

Vafsi has various strategies for disambiguation of A and P in these constructions:

- 1) The Agent is *always* Oblique-marked. Thus if there is only one Oblique it must be the A. This unambiguous interpretation, therefore, resolves all word order restrictions, even if there are no PAMs to help with disambiguation:

- (230) *tini kelj-e hágirætæ.*
 he:OB girl-DF₂ PVB-take:PT
 AGENT PATIENT PATIENT
 ‘He married a girl.’
- molla xáw-i yæ-værdae v-e.*
 P.N. sleep-OB PU-take:PT AUX:PT-3S₁
 AGENT
 ‘Sleep had overtaken Mullah.’

- 2) If there is at least one PAM (Set₁/Set₂) coindexing different persons or nouns of different number or gender, then word order is no longer an issue;
- 3) If the A and P are equal in type and sense and there are no PAMs (or they are also equal), OSV is the *default* word order that disambiguates A from P;
- 4) When Set₂ is fronted, it is commonly (but not obligatorily) hosted by the P;
- 5) Fronting of Set₂ to the Agent or over an Agent is not licensed. Thus in (227) above, the nouns hosting the Set₂ PAM can only be a P with SOV word order;

- (231) Double Oblique II (OSV)
leyle-y merde-y báé-diae-Ø.
 boy-OM man-OM PU-see:PT-3S₁
 ‘The man saw the boy.’
 ‘*The boy saw the man.’

- Double Oblique I (SOV) or II (OSV)
kelj-í merde-y báé-diae. (OSV)
merde-y kelj-í[=S] báé-diae. (SOV)
 man-OM girl-OF₂=3S₂ PU-see:PT
 ‘The man saw the girl.’
 ‘*The girl saw the man.’

While SOV (Double Oblique I) is preferred, if the A and P are equal (both M, F, or animate PL, or identical 3SG/PL pronouns), Sets_{1&2} cannot indicate which noun is co-indexed. Or if Set₂ is unfronted or deleted, the typical Set₂ host (P) is no longer the disambiguator. Only the *default* OSV word order can distinguish A from P.

(232) Double Oblique II (OSV)

<i>kærg-án=oan</i>	<i>luas-án</i>	<i>bé-værd-ende</i>
<i>kærg-án=oan</i>	<i>luas-án</i>	<i>b-ísan-værd-ende</i>
<i>kærg-án=oan</i>	<i>luas-án</i>	<i>bé-værdæ ↔</i>
<i>kærg-án=oan</i>	<i>luas-án</i>	<i>b-ísan-værdæ ↔</i>
chicken-OP=1P ₂	fox-OP	PU-3P ₂ -take:PT(-SET ₁)

'The foxes carried off our chickens.'

Double Oblique I (SOV)

<i>paša-y</i>	<i>dana-y</i> s	<i>bé-zæ</i>
---------------	---	--------------

Double Oblique II (OSV)

<i>dana-y</i>	<i>paša-y</i>	<i>bé-zæ-Ø</i>
<i>dana-y</i>	<i>paša-y</i>	<i>b-íz-zæ-Ø</i>
P.N.-OM	P.N.-OM	PU-3S ₂ -hit:PT-3SM ₁

'Pasha (m.) hit Dana (m.)'

7.2. Pronominal argument-marking and alignment in Tatic

Tatic pronominal alignment is generally mixed since different persons show different marking or Ø-marking. Thus Object-marking in pronouns can be more complex or less complex than nouns or may exhibit the same level of complexity, as, for example, in Vafsi pronouns with the same three alignments in the Past as nouns (§7.1.3.3), including the SOV/OSV issues. As for less morphological complexity, most of Tatic lacks a Direct-Oblique case distinction in the 2SG, 1PL, and 2PL.

Greater *morphological* complexity is seen in the cluster of Hezārrudi, Gandomābi, Nowkiāni, Kafteji which have innovated *two new cases* for P-marking in certain persons creating a four-way case distinction: 1) Present A/S, 2) Past A, 3) Pres. P, and 4) Past P. An additional possessive yields a total of five case forms in these persons (see §5.2.2). On the other hand, with separate tense-based splits in both Agent and Patient marking, *Alignment* here is Nominative-Accusative in the Present and Tripartite (Haig 2008: 188) in the past. In Tripartite alignment the agent of a transitive verb (A), the patient of a transitive verb (P), and the single argument of an intransitive verb (S) are each marked in different ways, viz., A is ergative, P is accusative and S is absolute. These varieties are in line with Haig's (2008: 189) Universal of Animacy-based splits in Iranian by which the alignment of SAP-pronouns is never higher on the Ergativity Continuum (p. 187) than that of nouns.

7.2.1. Tripartite and mixed pronominal inventories

Kafteji, the easternmost of the Tāromi group, shows all four cases in the 3SG, but the 1SG has merged the case forms for the Present P and Past P with the Genitive. See §5.2.2 for the Hezārrudi and Kafteji four-way distinction in 1st and 3rd person singular and also Yarshater (1970) for additional Hezārrudi examples of their use.

- (233) Kafteji
 Subject and Present Agent Present Patient (Nom-Acc Alignment)
 \boxed{a} $bə-dævást-\emptyset$. $tæ \quad \boxed{æja} \quad æ-m-vin-i?$
 he:DIR PU-run:PT-3SM₁ you he:ACC₁ PVB-DUR-See-2S₁
 'He ran.' 'Do you see him?'
 (234) Past Agent

- (234) Past Agent
 \boxed{ja} vend=es.
 he:OB see:PT=3s₂
 'He saw.'

- | | |
|-------------------------------------|--|
| Past Patient (Tripartite Alignment) | |
| <i>[jan]</i> | <i>bə-gé</i> ~ |
| he:ACC ₂ | PU-catch:PT ~ |
| 'I caught him/her.' | <i>jan=əm</i> <i>bə-gé.</i> |
| | he:ACC ₂ =1s ₂ PU-catch:PT |

7.2.2. Mixed pronominal typologies

Dikin Marāqeи pronouns show Present Nominative-Accusative (1SG, 3SG/PL), Present/Past Neutral (2SG/PL, 1PL), Past Ergative (3SG/PL) and Past Double Oblique (1SG) alignment.

- | | | |
|----------|---|---|
| (235) | PRESENT | PAST (Optional Set ₂ A agreement not shown) |
| Nom-Acc: | <u><i>æv-an</i></u> _{3SFD} <i>man</i> _{1OS} <i>m-in-</i>
<u><i>égan</i></u> _{3SF}
<u><i>æv-in</i></u> _{3PD} <i>æv-áé</i> _{3SOF} <i>m-in-</i>
<u><i>áén</i></u> _{3P}
<u><i>æz</i></u> _{1SD} <i>æv-i</i> _{3SOM} <i>m-in-éñ</i> _{1S} | <i>æv-áé</i> _{3SOF} <u><i>æv-in</i></u> _{3PD} <i>vínd-</i> Ergative
<i>in</i> _{3P}
<i>æv-on</i> _{3OP} <u><i>æv-an</i></u> _{3SFD}
<i>vínd-ian</i> _{3SF}
<i>man</i> _{1OS} <u><i>æv</i></u> _{3SDM} <i>vínd-</i>
<i>æ</i> _{3SM} |
| Neutral: | <u><i>šómæ</i></u> _{2P} <i>aemæ</i> _{1P} <i>m-in-ó</i> _{2P}

<i>aemæ</i> _{1P} <i>tó</i> _{2S} <i>m-in-óm</i> _{1P}
(Past only) \Rightarrow | <i>šómæ</i> _{2P} <i>aemæ</i> _{1P} <i>vínd-</i>
<i>iom</i> _{1P}
<i>aemæ</i> _{1P} <i>tó</i> _{2S} <i>vínd-iš</i> _{2S}
<i>æv-on</i> _{2OP} <i>man</i> _{1OS} <i>vínd-</i>
<i>in</i> _{1S} |
| | | Neutral Double Oblique |

'X sees/saw Y' (All examples SOV; Direct, **Oblique**, Neutral case is unmarked)

As in Dikin-Marāqeи, only 1SG. and 3SG/PL have Oblique case forms in Māsuleи and Asālemi, thus determining Double Oblique and Neutral alignment in Māsuleи with the addition of the Ergative in Asālemi. In Vafsi all persons except 2PL have Oblique forms. Only in Vafsi and Eštehārdi does the Direct case of the 1SG (*æz*) serve as Past Patient in a purely Ergative pattern:

- (236) Eštehārdi (ST: 149) Vafsi Ergative (Vafsi III, OSV)
 $\alpha\text{e}z=e\check{s}$ $b\ddot{\text{e}}-\text{r}\text{\textnormal{\textit{esta}}}=\text{y}\text{m}.$ $\alpha\text{e}z$ $t\text{a}\text{n}\text{i}$ $b\ddot{\text{e}}-\text{r}\text{\textnormal{\textit{esda}}}=\text{y}\text{m}$
I:DIR=3S₂ send:PT-1S₁ 'He sent me.' I:DIR he:OB send:PT-1S₁

7.2.3. The reassignment of the pronominal accusatives: Leriki

Northern Talyshi pronominal alignment is generally described as differing from the nominal alignment in the use of an innovated Accusative case for pronouns (except perhaps 3SG) to mark the P in all tenses. An analysis of Miller's texts—restricted to Leriki (Table 7A) for our present purposes—reveals incipient deviation from the use of the Accusative in the Past, while my own Leriki texts (Table 7B) show that this trend now seems to have been completely carried through in the 100 intervening years since Miller's texts. Additional modern materials may require some reassessment.

Table 7A: Miller (1930) texts: token count of pronoun marking

	Direct		Oblique		Common Case				Accusative Case					
Total:169	1SG	3SG	1SG	3SG	2SG	1PL	2PL	3PL	1SG	2SG	3SG*	1PL	2PL	3PL
Pres S/A	28	12			21	9	5	3						
Past S	6	3							5					
Past A			8	2	4				2					
Pres P			12		1				14	7	-	2		3
Past P		8		1	3			6	3	1				

Table 7B: Stilo (2002–2004) fieldwork texts: token count of pronoun marking

	Direct			Oblique			Common Case			Accusative Case				
Total:275	1SG	3SG	1SG	3SG	2SG	1PL	2PL	3PL	1SG	2SG	3SG*	1PL	2PL	3PL
Pres S/A	15	39	4		7	24	6	9						
Past S		4	28	21			16	1	2					
Past A			1	18	7	2	6		5					
Pres P				13						8	1	–	5	9
Past P		11	8		1	2		2		0	0	0	0	0

The small number of tokens for this survey could be problematic and comparison of additional Leriki as well as other dialects from Miller's 1930 texts (collected in 1902) vs. Stilo's 2000's Northern Talyshi texts will be needed to reach firmer conclusions—clearly a goal for future investigation. The points below on pronominal alignment are derived from Table 7B and thus show the more current (2002, 2004 fieldwork) status of Leriki:

The Direct case forms of the pronouns in the 2000's texts are used for:

- 1) Subject of transitives in the Present and all intransitive tenses (but see Point 7 below) as well as, very occasionally, the A of Past transitives (237b).

- (237) *az kitob-i o-bæ-gird-in-é=m.*
 I book-OB PVB-ALL-turn-CAUS-INF=AUX:1S₁
 'I will return the book.'

im diæ sæ b=aey.
 this look take-PT to=he:OB
 'It (the wolf) looked at him (boy).'

- 2) Patient in the Past system for 3rd singular pronouns:

- (238) *aev=ižon mayaré=dæ pæydo kard=e.*
 he=3P₂ cave=in manifest do:PT=AUX
 'They found her in a cave.'
im=im d=ištæ čas-í vindæ.
 this=1S₂ with=self eye-OB see:PT
 'I saw this with my own eyes.'

The Accusative case in the 2000's texts is used for:

- 3) the Patient in the *Present system*. SAP pronouns and 3rd plural show 23 tokens of the Accusative, and Ø tokens of Oblique or Common cases. The 3rd singular show 13 tokens of the Oblique case as the only alternative in the Present.

Caseless Pronouns: Neutral Alignment

- 4) The 2000's texts show 5 tokens of the Common case for the Patient in the *Past system* (not 1st or 3rd singular) and Ø tokens of the Accusative. Since there is no Oblique case in these persons, the A, T (i. e. the theme of ditransitive constructions), and P are all unmarked.

- (240) Agent
 $b=\ddot{i}$ $toyfæ=ku=mon$ æmæ $mæs-\alpha$ be.
 to=this tribe=from=1_{P₂} we hear:PT-PPL AUX:PT
 'We had heard about this tribe.'

Patient

*æmæ=žon yand=e.
we=3P₂ send:PT=AUX
'They sent us.'*

The Oblique case in my own texts collected in the field is used for:

- 5) 3SG Patient in the Present System as the preferred form. (See Pt. 3 above.)
 6) Agent of Past transitives in the 1SG *mi* (18 tokens) and 3SG *ey* (7 tokens).

- (241) *miː xæbæ sæ=y ki...
 I:OB news take:PT=AUX SUB
 'I asked them whether...'
 æy=æn vot=iš=e bæ taker-i.
 he:OB=also say:PT=3S₂=AUX to P.N.:OB
 'And he told Taher.'*

- 7) 1SG Patient in the Past System as the preferred form. Thus as with Dikin-Marāqeи and Asālemi (§7.2.2), N. Talyshi has a Double Oblique in the 1sg:

- (242) *iim-i* *mii* *bárd=iš=e.*
 this-OB I-OB take:PT=3S₂=AUX
 'He took me (there).'

iim-i *min=iž* *biabru* *ka=y.*
 this-OB I-OB=3S₂ disgraced do:PT=AUX
 'He disgraced me.'

- 8) The tokens of 1SG S-marking in the Present vs. the Past *intransitives* show almost mirror-image frequencies:

- (243) Present: Direct (78.9 %) Oblique (21.1 %)
- | | | |
|-----------------------|-----------|-------------------------------|
| <i>az=im.</i> | <i>mī</i> | <i>noxaš=im</i> |
| I=COP:1S ₁ | I:OB | sick=COP:1S ₁ EMOT |
| ‘It’s me.’ | | ‘Oh, I’m sick.’ |
- (244) Past: Direct (16 %) Oblique (84 %)
- | | | |
|--------------------------------|----------------|-------------------------|
| <i>az aeyil b-im.</i> | <i>mīn=aen</i> | <i>děštim.</i> |
| I child COP:PT-1S ₁ | I:OB=also | rise:PT-1S ₁ |
| ‘I was a kid.’ | | ‘I also got up.’ |

Differences between Miller’s 1902 texts and Stilo 2000’s texts:

MILLER (1930, COLLECTED IN 1902) STILO (FIELDWORK: 2002, 2004)

1. The 1SG Subject is exclusively Direct in both Present and Past systems

	The 1SG Subject is predominantly Direct (78.9 %) in Present system
	The 1SG Subject is predominantly Oblique (84 %) in Past system
 2. The Accusative case is used in both Present and Past systems (except 3SG.)

Accusative predominates in the Present, while in the Past, other than 3SG, marking of P with the Oblique (8) or Common (9) case is preferred over the Accusative (4) case.	The Accusative is used exclusively in the Present system (except 3SG.)
--	--
-
- | | |
|--|--|
| The P in the Past system is exclusively Oblique- or Common case-marked; the Accusative seems to have been ousted in this context | |
|--|--|

7.3. Set₁/Set₂: pronominal person-agreement markers

Set₁ (Direct) and Set₂ (Oblique) PAMs have various functions (see Table 5J, §5.3), but not all languages display the full range. Languages with a tense-sensitive split in the PAMs generally use Set₂ to encode the A in the Past. Only a few languages use Set₁ to co-index the P in all persons in the Past. A few more only have the latter usage in the 3rd person. Still others do not license any Set₁ encoding of P in the Past.

In addition, Set₁ functions as an independent enclitic copula (see Be₁, §6.8) and Set₂ is an enclitic form of the pronominal possessive (§4.2, Pt. 14). Based on a subset of functions of Set₁ and Set₂, there are nine (A-J, skipping I) language types:

Set ₂ encodes Past A	+	+	+	+	+	+	+	+	+	-
Set ₂ encodes Pres. P	+	+	+	+	-	-	-	-	-	+
Set ₁ encodes Past P	+	(+)	[3]	[3]	[3]	[3]	-	-	-	-
Set ₂ = Pron. Possessor	+	+	+	-	+	-	(-) ²	+	+	
Set ₂ encodes Past P	-	-	-	-	-	-	-	-	-	+
TYPE	A	B ¹	C	D	E	F	G	H	J ³	

- 1: Set₁ encoding of P in the Past is significantly more obligatory in Type A than in B. In Types C – F, the Set₁ is optional in this function and is restricted to the 3rd person, marked [3] here.
- 2: Set₂ possessive enclitics exist in Northern Talyshi but have a highly restricted usage.
- 3: Type J has Nominative-Accusative alignment in nouns, pronouns, and PAMs in all tenses

The nine types:

A. Dikin Marāqeи: Set₂ Past A; Set₂ Pres P, Set₁ Past P (all persons); Set₂ Possessive

- (245) Set₂ Possessive P Encoding, Pres. Set₁ P Encoding, Past
čašm=em, *čašm=et* *dæ-m-fa-n=eš*. *vínd=im-ian*.
 eye=1s₂ eye=2s₂ PVB-DUR-throw-1s₁=3s₂ see:PT=1s₂,3SF₁.
 ‘my eye, your eye’ ‘I throw it.’ ‘I saw her.’

B. Vafsi: Set₂ Past A; Set₂ Pres P, Set₁ Past P (all persons); Set₂ Possessive

- (246) Set₂ Possessive P Encoding, Pres. Set₁ P Encoding, Past
čosták=es, *čosták=esan* *is-aer-vin-óm* *b-isan-die-ym*.
 shoe=3s₂ shoe=3P₂ [3s₂]-DUR-see-1s₁ PU-3P₂-see:PT-[1s₁]
 ‘his/her, their shoes’ ‘I see him/her’ ‘They saw me’

C. Khoynarudi: Set₂ past A; Set₂ Pres P, Set₁ Past P (sparingly, restricted to 3SG/PL), Set₂ Possessive; (all: YSFN)

- (247) Set₂ Possessive P Encoding, Pres. Set₁ P Encoding, Past
baju=m, *zuré=r* *b-im-kešt-an!* *em-e=šun* *kešt-ind.*.
 sister=1s₂ son=2s₂ PU-[1s₂]-kill-2P₁ this-DP=3P₂ kill-[3P₁]
 ‘my sister, your son’ ‘Kill me!’ ‘They killed these ones.’

D. Nowkiāni: Set₂ past A; Set₂ Pres P; Set₁ Past P (restricted); no Set₂ Possessive; Patient Encoding, Present System: Set₁ A (suffixed)/Set₂ P (encliticized)

- (248) P Encoding, Pres. Set₁ P Encoding, Past
Ø-xeruš-ém=ešan. *béz-i-æ* *i-gæt=em-a.*
 DUR-sell-1s₁=3P₂ goat-INDF-DF PVB-take:PT=1s₂-[3SF₁]
 ‘I sell them.’ (YSFN) I bought a goat (f.).’ (YSFN)

E. Asälemi: Set₂ past A; no Set₂ Pres P; Set₁ Past P (restricted to 3SG/PL); Set₂ Possessive

- | | | |
|---|-------------------------|---------------------------------------|
| (249) | Set ₁ | Past P Encoding, |
| | Set ₂ | Possessive Past A marking omitted |
| šæví=m, zué=r | mən | æ merd-e vind-in. |
| shirt=1s ₂ son=2s ₂ | I:OB | that man-DP see:PT-[3P ₁] |
| ‘my shirt, your son’ | ‘I saw those men.’ | |
| | Set ₂ | A fronted |
| tə | æy-é=r | vind-in? |
| you | that-DP=2s ₂ | see:PT-[3P ₁] |
| ‘Did you see them?’ | | |

F. Hezārrudi: Set₂ past A; no Set₂ Pres P; Set₁ Past P (restricted to 3SG/PL); no Set₂ Possessive

- (250) Set₂, A unfronted/Set₁, P(YT: 458)

PU-put:PT=1S₂ [3P₁] storage.room
'I put them in the storage room.'

Past A marking omitted/Set₁ P (YT: 458)
aed-orund-ende otaq.
 PVB-throw:PT-[3P]_[...] room
 'X threw them into the room.'

G. Leriki: Set₂ A Past; no Set₂ Pres P; no Set₁ Past P; (Set₂ Possessive very marginal);

Past System: Set2 A (encliticized); Patient Encoding not licensed in any tense

- (251) *še-m* *mol-on=im* *dušæ=y.*
 go:PT-1s₁ cattle-PL=1s₂ milk:PT=AUX
 'I went (and) milked the animals.'

<i>omæ</i>	<i>maštævo</i>	<i>niyæ=m</i>	<i>o-šánd=e.</i>
come:PPL	morning	churn-1s ₂	PVB-shake-AUX
'The next morning I churned (it).'			

H. Sagzābādi: Set₂ A Past; no Set₂ Pres P; no Set₁ Past P; Set₂ Possessive

- (252) Possessive Enclitics (AA2: 103)
 $x^w ačær=i$, $dæst=eš$
sister=2s₂ hand=3s₂
‘your sister (Obliq.), his hand’

Past system: Set₂ = A (AA2: 103)
 $á̄sb-e=šun$ $dá̄-bæst$ dar .
horse-DP=3P₂ PVB-tie:PT tree
‘They tied the horses up to a tree.’

J. Razajerdi: no Set₂ past A; Set₂ Pres P; no Set₁ Past P; Set₂ Possessive; Set₂ Past P

- (253) Set₂ Possessive P Encoding, Pres. P Encoding, Past
bil=em, *sær=et* (not available) *bil* *bijæ* *zí-em=eš.*
 spade=1s₂ son=2s₂ spade with hit:PT-1s₁=3s₂ 'I hit it with the spade.'

7.4. Leftward mobility of person-agreement markers (Set₁/Set₂)

Except for Tatoid, Oskulak Talyshi, and Razajerdi, all under heavy Caspian and/or Persian influence, Fronting of Set₂ A PAMs is universally preferred but not obligatory in Tatic. By contrast, Set₁ clitic fronting is an isogloss in a restricted area including N. Tati, N. and C. Talyshi, as well as Armenian, Neo-Aramaic, and Udi (Stilo 2008a).

7.4.1. Fronting of Set₂ agent throughout the Tatic family

Set₂ clitics encoding the Past A may front to the immediately preverbal word ("One-Place Fronting"), most often the P, or to other elements farther to the left ("Distance Fronting"). They may also optionally remain on the verb, or in the right context be completely deleted. The examples below show this Set₂ flexibility in Tatic:

- (254) Talyshi: Leriki

Unfronted

hækim *vard-ðei=m=e.*
 doctor bring:PPL=1s₂=AUX
 'I have brought a doctor.'

One-place fronting (to object)

xariji *dil=aedæ* *gæymæti* *sæ=y?*
 foreign language=in grade=2s₂ take:PT=AUX
 'Have you gotten your grades in "foreign language" (yet)?'

- (255) One-place fronting (to non-object)

i *birdæm=dæi* *gæt=e* *mi.*
 one moment=in=3s₂ take:PT=AUX I:OB
 'It got me all of a sudden.'

Distance fronting (to object)

pulis *čey* *dast=o* *sæ=y.*
 money=3s₂ his hand=from take:PT=AUX
 'He took the money out of his hand.'

- (256) Distance fronting (to non-object)

bævædæ ⁺*mon* *gilæ-y hæyæjan dæ-voni*
 then-1P₂ UNC-INDF trouble PVB:pass:PT
 ‘Then we had (lit: passed) some trouble.’

Set₂ deleted

mī mæktub nīvīšt-é be.
 I:OB letter write:PPL AUX:PT
 ‘I had written a letter.’

In Asālemi Set₂ “... attaches to the left-most element in the phonologically integral verb phrase” (Paul 2011: 137), i. e., it floats leftward via a Wackernagel rule to a variety of initial-position elements. Set₂ may also optionally remain on the verb or be deleted:

- (257) Distance fronting: Object in 1
- st
- position

nælætí=m xəlik-i=næ kəšt=æ.
 snake=1s₂ shovel-OB=with kill:PT=AUX

Non-object in 1st position

xəlik-i=næ=m nælætí kəšt=æ.
 shovel-OB=with=1s₂ snake kill:PT=AUX
 ‘I killed the snake with a shovel.’

- (258) Unfronted

av-í dəlæ=ka per-aknást=əm=æ.
 water-OB in=KA PVB-pull:PT=1s₂=AUX
 ‘I pulled (it) up out of the water.’

Set₂ deleted

mən æ merd vind=æ.
 I:OB that man see:PT=AUX
 ‘I saw that man.’

- (259) Khoynarudi (YSFN)

Unfronted

sæmæd xan-i čest-é nur ⁺*čun* *=æ.*
 P.N. P.N.-OB shoe-DP put:PT=3P₂=AUX
 ‘They put out Samad Khan’s shoes.’

One-place fronting (to object)

šeme=ru korrey ⁺*em dur=æ.*
 you:PL=for trouble=1s₂ give:PT=AUX
 ‘I gave you trouble.’

- (260) Khoynarudi (YSFN)

Set₂ deleted

sæmæd xan-i vut=e ke...
 P.N. P.N.-OB say=AUX SUB
 'Samad Khan said that...'

- (261) Sagzābādi

Unfronted

num=eš u-jiræt [§sun] bæ'diæzzæman.
 name=3s₂ PVB-take:PT=3P₂ P.N.
 'They named him Ba'di-aëzzaman.' (AA2: 103)

One-place fronting (to object)

bæhram[=eš] be-zí zæmin.
 P.N.=3s₂ PU-hit:PT ground
 'He knocked Bahram to the ground.' (AA2: 115)

- (262) Distance fronting (to object)

teti[=em] eštæ-š owérd-e.
 girl=1s₂ you-also? bring:PT-PERF?
 'I have brought you the girl.' (AA2: 115)

Set₂ deleted

bæ'diæzzæman hučči né-va.
 P.N. nothing NEG-say:PT
 'B. didn't say anything.' (AA2: 107)

- (263) Nowkiāni (YSFN)

One-place fronting (to object) One-place fronting (to non-object)

tefæng[=es] u-gæ. *i dæfæ[=š] be-vint...*
 rifle=3s₂ PVB-take:PT one time=3s₂ PU-see:PT
 'He picked up his rifle.' 'Once he saw (that)...'

- (264) Nowkiāni (YSFN)

Distance fronting (to object)

bælkæ číz-i[=r] mez i-gæ.
 perhaps thing-INDF=2s₂ wage PVB-take:PT
 'Perhaps you got something for wages.'

Set₂ deleted

čo ma-r æjo be-va.
 his mother-OB he:OBJ PU-say:PT
 'His mother said to him,...'

(265) Vafsi

Unfronted

b-imresda koleng b-ísan-ard
 PU-1S₂-send:PT pick PU-3P₂-bring:PT
 'I sent them to bring a pick.'

One-place fronting (to object)

yey aynǣs b-ard.
 one mirror=3S₂ PU-bring:PT
 'She brought (him) a mirror.'

(266) Vafsi

One-place fronting (to non-object)

parama ke æz bén̄-es úgiræ.
 a.few EMF from root=3S₂ PVB-take:PT
 '(The flood) even picked up a few(trees) by the roots.'

Distance fronting (to object)

dàre si-é̄s æz ben o.bær árd
 tree-EZ apple-OF=3S₂ from root out brought
 'It ripped apple trees out by the roots.'

(267) Vafsi

Distance fronting (to non-object) Set₂ deleted

hæzírī-m tani bǣ-diæ. zarru-an injuri yá-watt-e.
 yesterday=1S₂ he:OB PU-see:PT child-OP thus PU-say:PT-PERF
 'I saw him yesterday.' 'The guys have said this (lit: thus).'

Set₂ fronting placement may help to identify the P in the Double Oblique I (§7.1.3.3).

7.4.2. Clitic restrictions in Vafsi

Set₂ clitic stacking is not allowed in Vafsi. If the only available host for the clitic already has a Set₂ as possessive, the second Set₂ will remain on the verb. Set₂ may combine, however, with other types of enclitics, e. g., =iz 'also.'

(268) *tin-an=iz gá=san b-ísañ-košd-æ.*
 this-PL=also cow-3P₂ PU-3S₂-kill:PT-PPL
 'They slaughtered their cows, too.'

(269) Unfronted (optional)

yeki ke=dæ b-íš-aesdæ.
 one house=in PU-3S₂-leave:PT
 'He left one (of them) at home.'

Fronting restricted: clitic stacking

<i>šæ:r=om</i>	<i>vel=on</i>	<i>kærd-e.</i>
city=1s ₂	loose-1s ₁	do:PT-PERF
'I have abandoned my country.'		

See also Experiencer verbs (§6.7) for similar examples of Set₂ fronting with 'to want'.

7.4.3. Set₁ fronting

Set₁ clitic fronting in West Iranian is restricted to Northern Tati, Northern and Central Talyshi, occurring only in analytic tenses: Present, Future, Progressives, Perfects, and in N. Tati in the Imperfect as well. This fronting is triggered by information structure by which the Set₁ encliticizes to the stressed focused element of the sentence, including even the Subject. Set₁ fronting is an areal feature also seen in Armenian, Neo-Aramaic, and Udi and its triggers are basically the same for both the Iranian and non-Iranian languages mentioned here. Since Set₁ clitics are also the AUX or copula, this fronting phenomenon also pertains to both Present and Past AUX forms:

(270) Leriki

Present		Future (see also ex. 45)
<i>čič=e</i>	<i>vót-dæ?</i>	<i>či</i> <i>kovra=ž</i> <i>b-æ-va?</i>
what=AUX:3S ₁	say-LOC	from where=AUX:2S ₁ FUT-DUR-bring
'What is he saying?' Where will you bring (it) from?'		

(271) Leriki

Past Progressive	
<i>tí</i> <i>čič=b-iž</i> <i>vot-é-dæ?</i>	
you what=AUX:PT-2S ₁ say-INF-LOC	
'What were you saying?'	

Past Perfect

<i>tožæ=b-im</i>	<i>az</i>	<i>omæ</i>	<i>kæ.</i>
new=AUX:PT-1S ₁	I	come:PT	house
'I had just come home.'			

(272) Viznei

Present		Past Progressive
<i>tæ</i> <i>čič=iš</i> <i>kárd-æ?</i>		<i>tæ</i> <i>čič=b-iš</i> <i>kard-æ?</i>
you what=AUX:2S ₁	do-LOC	you what=AUX:PT-2S ₁ do-LOC
'What are you doing?' (YSFN) 'What were you doing?' (YSFN)		

- (273) Asālemi

Present

æz æ otaq-i=ka=m b-æ-xət.
 I that room-OB=KA=AUX:1S₂ PR-DUR-sleep:PT
 'I sleep in that room.'

- (274) Keringāni

Present

Present Perfect

hæsæn=en pü'l get. nun=en=en pæt-a.
 P.N.=AUX:3S₁ money take:PT bread=1S₂=AUX:3S₁ bake:PT-PERF
 'Hassan gets money.' 'I have baked bread' (YSFN)

7.4.4. Simultaneous fronting of Set₁ and Set₂

Leriki transitive verbs in the Perfect tenses allow for both Set₁ and Set₂ to front simultaneously. Since the triggers for fronting are different for each set, they may front *together* or *separately* to different items (see also Keringāni 274, above):

- (275)
- aeyo ko kam=im=e kard=æ.*
- *

there work little=1S₂=AUX:3S₁ do=AUX
 'I worked there very little.'

*(scrambled version of: *kam ko=m=e kard-æ* 'worked very little')

idaræ=dæ=m=e ko kard=æ.
 office=in=1S₂=AUX:PERF work do=AUX
 'I (have) worked in an office.'

- (276) Past Perf., Fronted together Past Perf., Fronted Separately

tarix iminji=m=be niivištæ. *tarix=im iminji=be niivištæ.*
 history first-1S₂=AUX:PT wrote history-1S₂ first=AUX:PT wrote
 'I had registered (lit: written) history as (my) major (lit: first).'

7.5. Ditransitives: recipients and addressees

Most Caspian and Tatic languages exhibit Neutral Ditransitive typology (T=P=R): the Theme (T), Recipient (R) of Ditransitive verbs, and Patient (P) of monotransitives are identically flagged with =ra in Caspian, Harzani, Razajerdi and Tatoid and for the most part with the Oblique case in the rest of Tatic. This strategy is also called the Double Object construction. This typology does not apply to non-salient Ps and Ts which are unmarked while the R is always flagged with the respective strategies mentioned, producing an Indirective typology, i. e., P=T≠R (see also Stilo 2010). There are also other minor adpositional patterns that flag the R, yielding Indirective typology. N. Talyshi of the Republic of Azerbaijan uses the latter typology exclusively.

Depending on the language, Addressees may be flagged in the same way as Rs although frequency counts may reveal that some languages show slight differences in postverbal vs. preverbal word orders of Recipient and Addressees, respectively. See §9.2 for a discussion of word order issues for Recipients and Addressees.

(277) Lāhijāni

*mu bút-furt-əm digár=e**.
I PU-SELL:PT-1S other-INDF:RA

'I sold (it) to another (i. e., someone else).' *(*digár=e < digár-i=a*)
(NJ: 61)

mu ketab=a un=a há-da-m
I book=RA he=RA PVB-give:PT-1S
'I gave him the book.'

Addressees

(278) *be:tær bu-Ø t-æ b-utt-ə bu-m.*

better COP:PT-3S you-RA PU-say:PT AUX:PT-1S
'It would have been better had I told you.'

tu k-e b-ut-i?*
you who-RA PU-say:PT-2S
'Who did you tell?' (NJ: 55)
(**ki* 'who' + =a > *ke*)

(279) Sāravi

Recipient

sabsere dær-e ve=re ešnasek de-n-e.
landlord BE₄3S he=RA hint give-PR-3S
'The landlord is giving him a hint.' (SY: 60)

Addressee

qeziye=re še hæmsaye=re bæ-ut-e.
issue=RA self neighbor=RA PU-said-3s
'He told the situation to his neighbor.' (SY: 62)

(280) Harzani

Recipient

<i>en jöb=e te=re ü-müt=mæ.</i>	<i>merd=e öt=mæ.</i>
this word=RA I=RA PVB-teach=1s ₂	man=RA say:PT=1s ₂
'I taught you this word.' (MH: 466) 'I said to the man.' (AAK: 78)	

Northern Talyshi flags the Recipient and the Addressee exclusively with the adposition *bæ* ‘to’, while Asālemi appears to use both Oblique case and *bæ* strategies, and Māsulei seems to use only the case-marking strategy, as do most Tatic varieties.

(281) Leriki

Recipients

b-æ-do=n *æy* *bæ* *mī?*
 FUT-DUR-give=3_{P₁} that:OB to I:OB
 ‘Will they give that to me?’

Addressees

æv-on *bæ* *mī* *nī-bæ-vot-e=n.*
 that-PL to I:OB NEG-FUT-DUR-say-INF=3_{P₁}
 ‘They won’t tell me.’

(282) Asālemi

sif-i *b-æ-da=m* *æ:mæd-i.*
 apple-OB FUT-DUR-give=1_{S₁} p.N.-OB
 ‘I’ll give Ahmad the apple.’

ræis-i *bæ* *mən* *vat-æ* *ki...*
 chief-OB to I:OB say:PT-AUX SUB
 ‘The chief told me that...’ (Paul: 191)

(283) Māsulei

pul-e *a-do=š-æ* *mən.*
 money-OB PVB-give:PT=3_{S₂} I:OB
 ‘He gave me the money.’

æ *mærdæk-é* *əčə=y* *vat=æ?*
 that man-OB what=2_{S₁} say=AUX
 ‘What did you say to that man?’

7.5.1. Vafsi ditransitives

Vafsi has three strategies for ditransitives (see also Stilo 2010 for a fuller discussion of ditransitives in Vafsi). Of these, the two major constructions are:

- the Double Object Construction (DOC) in which the R is Oblique-case marked;
- the Indirect Object Construction (IOC) in which the R is flagged with adpositions.

The DOC, by far the most common ditransitive construction in Vafsi, typically (but not obligatorily) places the R argument in postverbal position:

- (284) *in ke t-æsbir-ám esdæ.*
 this house DUR-entrust-1_{P₁} you.OB
 'We are turning this house over to you.' (VF: B5,13)
- púl=om há-da hæsén-i.*
 money=1s₂ PVB-gave P.N.-OB
 'I gave (the) money to Hassan.'

As for the IOC and adpositional flagging, the preposition *o* occurs in the Vafs dialect mostly with the verb 'give' as an occasional alternate to the DOC:

- (285) *kellæ dokán-i=s háda o merde-y.*
 key shop-OM=3s₂ PVB-gave to man-OM
 'He gave the man the key to the shop.'
- in leyle-y æt-æsbir-om o esdæ.*
 this boy-OM DUR-entrust-1s₁ to you:OB
 'I am entrusting this boy to you.'

In most languages of western Iran, a very limited set of verbs, esp. 'write' and 'send', require the adpositional flag 'for' to mark the Recipient rather than an allative flag:

- (286) Vafsi
ketab=om bæ-restá=ve tani=ra. nawé=m bæ-nivisdaæ esdæ=ra.
 book=1s₂ PU-sent=AUX he:OB=for book=1s₂ PU-sent=AUX he:OB=for
 'I had sent him a book.' 'I wrote you a letter.'
- (287) Sāravi (SH: 157)
æt-ta kavez še zena vesse bæ-nvišt-eme
 1-UNC letter self wife for PU-write:PT-1s
 'I wrote my wife a letter.'

The Addressee in Vafsi is most commonly flagged by the preposition *dæ* or sometimes by the Oblique case as in the DOC. The Addressee is generally preverbal.

- (288) *dæ tawan hič=es ná-wattæ.*
 to we.OB nothing=3s₂ NEG-said
 'He didn't say anything to us.'
- e-yz=aem in ær-vaz-om esdæ.*
 I.DIR=also=also this DUR-say-1s₁ you.OB
 'And I will tell you this.' (B12.5)

The third, or minor, Ditransitive pattern of Vafsi is the PAM Ditransitive Construction, an *indexing* strategy that encodes a pronominal R with Set₂ PAMs directly on the verb. With 'give', the Set₂ clitic is *obligatorily* fronted to the Theme. Examples:

- (289) *[is]-ær-vaz-om.* *ær-goa* *ke* *nešanæ-[s]* *bæ-d-a.*
 3S₂-DUR-say-1S₁ DUR-want:PT SUB address=3S₂ PU-give-2P₁
 'I'll tell him/her.' 'You had to give him an address.'

7.6. Other dative-like usages

A few other minor patterns with a Dative sense take Caspian =ra marking or Tatic Oblique-case marking. They seem to be triggered by lexical items that require them:

- (290) Vafsi
*æ-čuæ** *esdæ* *tæmen* *bæ-ræs-an-e?*
 DUR-can you:OB I:OB PU-arrive-CAUS-1S₁
 'Can he bring you up to me (i. e., make you my equal?)' *(*æ-čuæ < at-suæ*)

Lāhijāni (NJ: 115)
pəriruz *šəmə=ra* *či* *bu-Ø?*
 day.b.yest you:PL=RA what COP:PT-3S
 'What was wrong with you the day before yesterday?'

- (291) Sāravi (SY: 58)
še *mar=e* *mæmnundar* *bi-ye.*
 self mother=RA grateful COP:PT-3S
 'She was grateful to her mother.'

Māsulei
tə *mən* *Ø-mon-e.*
 you I:OB PR-look.like-2s
 'You look like me.'

8. Subordination

Subordination types all share the feature of being introduced by the Universal Subordinator *ki/ke* (SUB) in their underlying structures. This survey of Subordination briefly covers Adverbial ('when, if, since'), Relative, Complement, and Purposive clauses.

8.1. Word order typology of subordinate clauses

As we shall see in §9, word order issues in the languages of this area indicate a rather mixed typology. Subordinate clause types also demonstrate an equally mixed

word order typology. They are generally left-branching, but right-branching in a significant number of major subordinate clause types. What Diessel (2001) shows for the word order of Persian adverbial clauses holds equally true for most, if not all, languages of western and central Iran: Conditionals predominately precede the main clause, Temporals generally do, Causals are mixed, and Purposive (to which I will add Complement and Relative) clauses always follow the head.

8.2. Adverbial clauses

Adverbial clauses have the following characteristics:

1. They are introduced by a subordinate conjunction, in which case the SUB may be retained or optionally deleted (exs. 292–303);
 2. Additionally, the temporal, causal, or conditional subordinate conjunction is commonly deleted. The SUB is then retained, and the semantics of the three types are all neutralized. With this deletion, if the enclitic SUB *ki/ke* is in initial position, it must be shifted to the right of the first constituent (exs. 304–313);
 3. If the adverbial clause is shifted to after the main clause and if the subordinate conjunction is deleted, the SUB is no longer initial and is not right-shifted;
 4. Initial adverbial clauses have a very distinctive intonational contour (also see pitch levels as discussed in §3.6), marked with (,,) in this chapter: all the usual word and sentence stresses and pitch of the whole clause are neutralized, the first syllable of the clause gets a high pitch, and the last syllable of the clause has an extra-long extension with the pitch consistently rising to extra-high.
 5. Right-shifted initial adverbial clauses have a different distinctive contour: all stresses are again neutralized, the pitch remains flat and at an extra-low level throughout the clause.
 6. The subordinate conjunction and the SUB may also be optionally deleted, in which case the only marker of subordination is intonation (exs. 314–318);

Examples marked with * below indicate that both the adverbial subordinate conjunction and the Universal Subordinator *ki/ke* are present.

8.2.1. Temporals

- (292) Lāhijāni
 váxti mu ún=a bə-dye-m,, u xott-í du-bu-Ø.
 when I he=RA PU-SEE:PAST-1S he sleep:INF BE₄-AUX:PT-3S
 'When I saw him, he was sleeping.'

- (293) Sāravi (SH: 151)

váeqti *æt* kæm-i gæt bæ-yy-e,, še mar-e
 when one little-INDF big PU-become:PT-3S self mother-REZ
qædr=e fæ:m-en-e.
 value=RA understand-PR-3S

‘When she gets a little older, she will appreciate her mother.’

- *(294) Leriki (MT: 114)

váxti=kí æ moy=š vind=e,...
 when=SUB that fish=3S₁ see=AUX
 ‘When he saw that fish,...’

- Kajali (KJ: 285)

váexti=ke jællad=eš be-hind,...
 when=SUB executioner=3S₂ PU-SEE:PT
 ‘When she saw the executioner,...’

- *(295) Vafsi

héni=ke vá-nde yeyláq-i séyl-æ b-áw-e,,
 when=SUB say:PT-3P₁ winter.camp-OM flood-DF PU-come:PT-3S₁
æz bæ-tærsa-yme.
 I PU-fear:PT-1S₁

‘When they said a flood came at the winter camp, I got scared.’

8.2.2. Conditional clauses

(for additional examples see also: Conditionals, §6.5.12)

- (296) Lāhijāni (LTXT)

agə bæ-xa-m futbol bazi bù-kun-əm,,
 if PU-want-1s soccer game PU-do-1s
 ‘If I want to play football,...’

age mu dubare bæ-kye-m,,...
 if I again PU-fall-1s
 ‘If I fall again,...’

(297) Sāravi

*æger te bē-xa-i,, xærim-be**,...
 if you PU-want-2s
 'If you want, I'll buy (it).' (< *xærin-me) (SH:143)

æge hič ja mimuni b-úr-em.
 if any place party PU-go-1s
 'If I go to a party anywhere, ...' (SY:58)

(298) Leriki

ægær merd-í ištæ nom=iž sipriš noa bu-Ø,,...
 if man-OB self name=3s₂ elder put AUX:SBJ-3s₁
 'If a man has called himself a wiseman,...'

*(299) Vafsi

áge ke æz pa=y æt-eyšer-ome,,...
 if SUB I foot=2s₂ DUR-squeeze-1s₁
 'If I squeeze your leg, ...'
ča æmiq-i ágæ bæ-zæn-ende,,...
 well deep-INDF if PU-hit-3P₁
 'If they strike a deep well, ...'

8.2.3. Causal clauses

There are two primary types of causal adverbial clauses. The *čon*, *čun*, *čunki* ‘since’ type usually precedes the main clause, while the ‘because’ (Persian: *bæráye ínke*, etc.) type generally follows. The latter type is discussed below in §8.3.

(300) Lāhijāni (NJ: 161)

čun nəhar šime məymon b-im,, nú-tonəs-im orə
 since lunch your:PL guest COP-1P NEG-can:PT-1P there
bá-š-im.
 PU-go-1P
 'Since we were your guest for lunch, we couldn't go there.'

(301) Sāravi (SH: 143)

čon xæssu=me, sere mund-em-be. (< *munn-en-me)*
 since tired=COP:1s₁ house-PR-1s stay-PR-1s
 'Since I'm tired, I'll stay home.'

- (302) Māsulei
 čon hæva gærmtær=a bo-Ø, mosafər hey
 since weather hotter=CS become:PT-3S₁ traveler keep.on
 aeræy kær-ə.
 sweat do-3S₁
 ‘Since it got hotter, the traveler keeps sweating.’
- (303) Vafsi
 čun næf'e xosusi man beyn=dæ ná-v-e,, tæmam
 since gain-EZ private in middle=in NEG-COP:PT-3S₁ all
 esteqbal=esan kærd
 welcome=3P₂ do:PT
 ‘Since there wasn't any private gain involved, they all welcomed it.’

8.2.4. Subordinate conjunction deleted, *ki* shifted

Temporals

- (304) Lāhijāni (LTXT)
 dišow ki hæmdiga=ra bə-d-im, ...
 last.night SUB each.other-RA PU-see:PT-1P
 ‘Last night when we saw each other,...’
- dər-ə=ji ki b-áma-m birun,,
 door-EZ=from SUB PU-come:PT-1S out
 ‘When I came out the door,...’
- (305) Sāravi (SY: 66)
 zena ke sere=je b-urd-e, ...
 woman SUB house-from PU-go:PT-3S
 ‘When the woman left the house,’
- (SY: 54)
 sob ke bæ-i-ye, ...
 morning SUB PU-become:PT-3S
 ‘When it became morning,’
- (306) Leriki (MT: 115)
 i xæbæ-y ki bæ podšo ræsæ-Ø podšo mæyus
 this news-? SUB to king arrive:PT-3S₁ king depressed
 be-Ø.
 become:PT-3S₁
 ‘When this news arrived to the king, the king got depressed.’

- (307) Hezārrudi (YT: 466)

*ævvæl ruze ke æz xordæyal vi-em,, čemen dad
 first day SUB I child COP:PT-1S₁ my father
 be-mbæ-Ø.*

PU-die:PT-3S₁

‘In the beginning when I was a kid, my father died.’

- (308) Vafsi

*næzdike əwdaní ke váe-ræsa-nde,, vel=esan kærd.
 near-EZ village SUB PU-arrive:PT-3P₁ loose=3P₂ do:PT
 ‘When they got near the village, they released (it).’*

Conditional Clauses

- (309) Lāhijāni (NJ: 98)

*šəmə ki išon-ə ama qa:r n-in, pəs čərə
 you:PL SUB they-REZ with angry NEG:COP-2P then why
 išon-ə xonə n-a-n-in?
 they-REZ house NEG-come-PR-2P
 ‘If you are not mad at them, then why won’t you come to their house?’*

Causal Adverbial Clauses

- (310) Lāhijāni (NJ: 177)

*tu ki riš=a təraš-ən-i, čera ti sibil=a
 you SUB beard=RA shave-PR-2s why your mustache=RA
 nə-zən-i?
 NEG-hit-2s*

‘Since you shave your beard, why don’t you shave your moustache?’

- (311) Vafsi

*in ke soan há-da tæmen,, æz is-æt-ar-om
 this SUB you:PL PVB-give:PT I:OB I 3S₂-DUR-bring-1S₁
 soan goftošunúd hákær-a.*

you:PL conversation PVB-do-2P₁

‘Since you gave me this, I will bring him so you can converse (with him).’

The rightward shifting of the SUB may also occur more than one place to the right:

- (312) Khoynarudi (YSFN)

*ferqæći-ün em=čün šahar ke estéræ,
 P.N.-OP this=3P₂ city SUB take:PT
 ‘When the Ferqa people took this city,...’*

Sagzābādi (AA2: 107)

lælǣ jem gæp=eš ču be-šnowost
 nanny this:OB speech=3s₂ SUB PU-hear:PT
 'When the nanny heard this,...'

(313) Vafsi

par awan ke væ-ss-iam kongræ...
 last.year we SUB PU-go:PT-1P₁ congress
 'Last year when we went to the congress,...'

8.2.5. Subordinate conjunction and ki deleted (temporal, conditional, causal)

(314) Lāhijāni

šom xord-ə=də b-im,, nahid b-ámə-Ø ame xonə.
 dinner eat-INF=BE₄ AUX:PT-1P P.N. PU-come:PT-3s our house
 'When/while we were eating dinner, Nahid came to our house.' (NJ,134)

(315) Sāravi (SY: 58)

e-tta tim karess-ene,, sæt-ta tim æmel
 one-UNC seed plant:PT-3P 100-UNC seed produced
Ø-yæmu-Ø.
 IMF-come:PT-3s
 'If they planted one seed, a hundred seeds would grow.'

(316) Leriki

æy bř-vind-o-ž, bæ-zn-e=ž?
 he:OB SBJ-see-OPT-2S₁ FUT-know-INF=2S₁
 'If you see him, will you recognize him?'

(317) Vafsi

in pæs=dæ tt-á-nde,, æz rótt=i m-eyz.
 this time=in DUR-come-3P₁ from place=2S₂ NEG:IMPER-rise
 'This time when they come, you don't get up from your spot.'

(318) Vafsi

suraw bř-zen-ende,, áwæ obær ætt-à-Ø.
 P.N. PU-hit-3P₁ water to-out DUR-come-3S₁
 'If they drill in Shuraw, the water will come out.' (i. e., they'll strike water)

8.2.6. Right-shifted adverbial clauses

- (319) Lāhijāni
- | | | | | |
|--------------|------------------|-------------|------------|-----------------|
| <i>mu</i> | <i>n-ámə-na</i> | <i>bu-m</i> | <i>agə</i> | <i>bú-donsə</i> |
| I | NEG-COME:PT-COND | AUX:PT-1S | if | PU-KNOW:PT |
| <i>b-i-m</i> | <i>itorə=y.</i> | | | |
- AUX:PT-IRR-1S this.way=COP:3S
 ‘I wouldn’t have come if I had known that it was this way.’
- (320) Lāhijāni (NJ: 71)
- | | | | | | |
|-----------|--------------|-------------|-----------|----------|-----------------|
| <i>mu</i> | <i>bu-šo</i> | <i>bu-m</i> | <i>ki</i> | <i>u</i> | <i>b-áma-Ø.</i> |
| I | PU-GO:PT | AUX:PT-1S | SUB | he | PU-COME:PT-3S |
- ‘I had (already) left when he came.’
- Leriki
- | | | | |
|-----------------------------------|--------------|-----------|---------------------------------------|
| <i>bærk=imon</i> | <i>še-dæ</i> | <i>ki</i> | <i>társ-dæ=mon</i> |
| strong=1 _{P₁} | go:INF-PR | SUB | fear-PR-PR=1 _{P₁} |
- ‘We are walking fast because we are scared.’
- (321) Vafsi
- | | | | | | | | |
|----------------|-------------|-------------------------|-----------|-----------------------|----------|--------------|------------|
| <i>molla-y</i> | <i>bawæ</i> | <i>áe-veš-e</i> | <i>ke</i> | <i>hoqqæ=s</i> | <i>o</i> | <i>tawan</i> | <i>zæ.</i> |
| P.N.-OM | father | PU-burn-3S ₁ | SUB | trick=3S ₂ | to | we:OB | HIT:PT |
- ‘May Molla’s father burn (in hell) because he has played a trick on us.’
- 8.3. Right-branching subordinate clauses (main-sub)
- Subordinate clauses whose expected word order is MAIN-SUB primarily consist of three types: Causal II, Complement, and Purposive clauses. Relative clauses also follow the head noun but may precede (as one alternate) in N. Talyshi (338–39).
- (322) Leriki (MT: 100)
- Causal II
- | | | | | | | |
|----------------------|------------------|----------------|-----------|-----------|-----------|-----------|
| <i>æv=iš</i> | <i>ní-znæ=e</i> | <i>sápinae</i> | <i>æy</i> | <i>ki</i> | <i>ve</i> | <i>bæ</i> |
| s/he=3S ₂ | NEG-KNOW=AUX | because.of | that:OB | SUB | very | to |
| <i>yandi</i> | <i>áe-š-i-n.</i> | | | | | |
- each.other DUR-GO-IMF-3P₁
 ‘He didn’t recognize her because they looked a lot like each other.’

8.3.1. Complement clauses

(323) Lāhijāni

ama bá-d-im ki mi pier=əm b-áma-Ø.
 we PU-SEE:PT-1P SUB my father=also PU-COME:PT-3S
 ‘We saw that my father came in.’

(324) Leriki

vínd=iž=e ki im še-Ø jali.
 see=3S₂=AUX SUB this go:PT-3S alone
 ‘(The wolf) saw that this one went off (from the pack) alone.’

ve say be ki tř omáč=ž.
 very healthy be:IMPER SUB you come:PT=2S₁
 ‘Thank you very much for having come. (lit: that you have come).’

(325) Khoynarudi (YSFN)

vin=čun=æ¹ asb-iin bær an=æ² qefel.
 see:PT=3P₂=AUX horse-OP door BE_{DEM}=COP:3S locked
 !(*< vind=šun=æ*)

²(for copula type see at §6.5.8.1)

‘They saw that the horses’ (stable) door is locked.’

(326) Vafsi

esdæ ná-watt-æ ke an čiz b-e.
 you:OB NEG-SAY:PT-PPL SUB that what AUX:PT-3S₁
 ‘You didn’t say what that was.’

b-ísti ke náé-r-buæ-Ø.
 PU-3S₂-SEE:PT SUB NEG-BE:3S₁
 ‘He saw that it isn’t possible.’

8.3.2. Purposive clauses

Purposive clauses are usually introduced either by a subordinate conjunction that also means ‘until’ (*ta*, etc.) or by the SUB *ke/ki*. Both subordinators may also be deleted. The verb in the Purposive subordinate clause is always in the Present Subjunctive:

(327) Lāhijāni (NJ: 184)

*dariyuš i harf-on=a zen-ə ta b-əsyn-ə širin či
 P.N. this speech-PL=RA hit-3S until PU-SEE-3S P.N. what
 gu-n-əy.
 say-PR-3S*

‘Dariush is saying these things (words) to see what Shirin will say.’

- (328) Sāravi (SH: 150)

višter bæ-xos ta æsæbuni n-ævv-i.
 more PU-sleep until angry NEG-become-2s
 'Sleep more so you don't get angry.'

(SH: 143)

bi-yamu-Ø ke ve=re bæ-ver-e.
 PU-come;PT-3S SUB he=RA PU-take-3S
 'He came to take him.'

- (329) Leriki

čæpæ=mon ka ki gin ná-bu-n.
 fence=1P₂ do:PT SUB lost NEG-become-3s
 'We made a fence so they don't get lost.'

Sagzābādi (AA2: 105)

be-š-ime da jiæ b-ar-em.
 PU-go-1S₁ until she:OB PU-bring-1S₁
 'I went to bring her back.'

- (330) Vafsi

ær-go vel-an áw-æ há-de-y yáwa
 DUR-want flower-PL water-DF PVB-give-2S₁ until
né-melesen-u-æ.
 NEG-wither-PASS-3S₁
 'You have to water the flowers so they don't dry up (wither).'

- (331)
- æhmédi vasettáe=s kærdæ-ve ke awan ašdī*

P.N. mediator=3S₂ do:PT=AUX:PT SUB we peace
hà-kær-am.
 PVB-do-3P₁
 'Ahmad (had) acted as a mediator so that we would make up.'

8.3.3. Postposed relative clauses (see §8.4.1 for preposed)

- (332) Sāravi

un merdi=re ke vin-d-i me piyer=e.
 that man=RA SUB see-PR-2S my father=COP:3S
 'The man that you see is my father.'

Vafsi

kelje ke mákk=a w-è, indi=e.
 girl SUB lost=CS be:PT-3S here=COP:3S₁
 'The girl who got lost is here.'

Relative clauses modifying an object are often extraposed to the right, particularly if they are longer than the verb of the main clause (cf. a similar situation in German).

(333) Lāhijāni

xatire-i bu-Ø ki mu dašt-əm.
memory-RESTR COP:PT-3S SUB I have:PT-1S
'It was a reminiscence that I had.'

ja-ha-yi hanna ki mu nú-šo-m.
place-PL-RESTR BE₅ SUB I NEG-GO:PT-1S
'There are places I haven't gone to'

(334) Lāhijāni (NJ: 130)

mu xəyli čiz-on=a don-əm ki gut-ə mənne-m.
I many thing-PL=RA know-1s SUB say-INF can:NEG-1S
'I know a lot of things that I can't say.'

Since the subordinator *ke* is invariable, NPs that have a non-core argument role in the relative clause ('whose, with whom', etc.) are generally replaced by a resumptive pronoun, either full form or *Set₂*, encoded in the same role as the underlying NP:

(335) Vafsi

an merde-y ke nawé=s i-r-pærsa, ke ve-Ø?
that man-RESTR SUB name=3s₂ 2s₂-DUR-ask:PT who COP:PT-3s₁
(nawé=s ~ tani nawé)
(name=3s₂ his name
'Who was that man whose name you were asking about?'

8.4. Non-finite types of subordination in Leriki

In N. Talyshi temporal and purposive clauses have common non-finite alternatives built on the infinitive of the main verb and an adposition: temporals with the locative postposition =ædæ and purposives with the benefactive preposition *bo*, as in Leriki:

Temporal

(336) čoštæ hard-i=ædæ ov=i_m hárð=e.
breakfast eat-INF=LOC water=1s₂ eat:PT=AUX
While eating breakfast, I drank water."

tí væy b-i=ædæ, bárd=e?
you there be:INF=in take:PT=2s₂
'When you were there, did you take (him/her with you)?'

Purposive clauses

- (337) *penj dæyiþæ iyo mánð-imon bo hærækæt kard-e ištaen tosp-in-e.*
 five minute there stay-1_{P₁} for movement do-INF self
 warm-CAUS-INF
 'We stayed there for five minutes to move around (and) warm ourselves.'

8.4.1. Preposed relative clauses

Leriki has both postposed and preposed Relative clauses. The latter variant is a non-finite type based on the past participle with the same structure for core (S, A, P) or oblique roles (Possessive, Locative, etc.) of the deleted NP:

Core functions of deleted NP

- (338) *dī asp-i šæ kæs híst=e.*
 with horse-OB go:PPL person exist:=COP:3_{S₁}
 'There are people who go by horse.'
čimí bæ tī doæ kitob kónjo=ye?
 I:POSS to you give:PPL book where=COP:3_{S₁}
 'Where is the book I gave you?'

Peripheral functions of deleted NP

- (339) *ištī gæp žæ zoæ kív=e?*
 you:POSS speech hit:PPL boy where=COP:3_{S₁}
 'Where is the boy *you were talking to*?'
ko kard-æ vîræ=dæ tī tok-í žæ.
 work do-PPL place=in you current-OB hit:PT
 'The (electric) current struck you at the place *where you work*.'

9. Additional word order issues

Most Tatic and Caspian languages exhibit predominately left-branching typology but there are significant exceptions, e. g., right-branching Relative, Complement, and Purposive clauses, consistently postverbal Goals and Recipients, among others. In all of the languages in this area, esp. Caspian, we find many (and still on-going) syntactic copies from Persian, e. g., Set₂ Oblique enclitics (§6.7), various right-branching patterns in the NP (the *Ezafe*, prepositions, etc.). In addition, all of Tati and most of Talyshi are also in heavy contact with Azeri, which is predominately left-branching.

We have already discussed the left- and right-branching patterns in subordination (§8.1) above. In addition, languages such as Vafsi, located in transitional zones between the northerly left-branching languages and the southerly right-branching languages (see Stilo 2005), have highly mixed typologies in adpositions (§5.7.1), genitives (§5.4.3), and, to a lesser extent, adjectives. In the next section we will see further right-branching and mixed word order typologies of Patients/Objects, Goals, Recipients and Addressees.

9.1. Postverbal Objects

As attested throughout the examples given in this chapter, the basic word order of core arguments of the clause is SOV, but Object-Verb word order is not rigid or absolute. While postverbal objects are commonly heard in speech, they occur exclusively with definite nouns in these languages (see exs. 239, 255, 281). Frequency counts have not yet been conducted for most of these languages, but we can take the situation found in Miller's (1930) Leriki corpus, Lazard's (1979) Māsulei corpus and comparable segments of my own Lerik texts and of Amini's Tonekāboni texts in Table 9A as typical.

Table 9A: OV/VO word order

	Leriki:Miller	Leriki:Stilo FN	Māsulei:Lazard	Tonekāboni: Amini
Def OV	47	88.7 %	32	86.5 %
Def VO	6	11.3 %	5	13.5 %
Indef OV	11	100 %	34	100 %
Indef VO	0	0 %	0	0 %

In addition, it is interesting to see just how many non-core arguments are preverbal in a language such as Vafsi. While frequency counts for postverbal objects have not yet be done for Vafsi, the following table shows just how common other postverbal elements are.

Table 9B: Vafsi preverbal vs. postverbal non-core arguments

Argument	Addressee	Temporal	Ablative	Locative	Comit/Instr.	Bene-factive	Recipient	Goal
Preverbal	96.3 %	95.2 %	93.2 %	88.9 %	84.2 %	64.6 %	33.7 %	13.8 %
Postverbal	3.7 %	4.8 %	6.8 %	11.1 %	15.8 %	35.4 %	66.3 %	86.2 %

The most striking observation about this pull chain is the predominant postverbal nature of Goals vs. the predominant preverbal position of the cluster of Addressees, Temporals, Ablatives, Locatives, and Comitative/Instrumentals. Note that the Locative at 88.9% preverbal is almost the mirror-image of the Goal at 86.2% postverbal. It is my opinion that the Goals have initiated this pull chain and since Recipients often share many features with Goals cross-linguistically, as do Benefactives with Rs, these arguments followed in the chain as seen in the table. There is still a rather large gap between the three arguments on the right and all five others to the left in the table.

9.2. Goals, recipients, and addressees

As we see in the Table 9B for Vafsi, Goals are generally postverbal by near-universal, yet optional, preference throughout the area. In Borjian's Kordkheyli texts, for example, slightly less than 10% of Goals are in preverbal position. By contrast, I found no postverbal Locatives in this corpus. Note that Goals, except in Leriki, are generally unmarked by adpositions or case, i. e., are Ø-marked. Word order, while not obligatory, is mostly the distinguishing factor.

- (340) Lāhijāni
- | | | | | | |
|---------------------|----------------|--------------|-------------------------------|------------------|--------------------|
| <i>ba:d</i> | <i>bú-šo-m</i> | <i>xonə.</i> | <i>m-a</i> | <i>bá-vərd-I</i> | <i>bimarestan.</i> |
| then | PU-go:PT-1s | house | I-RA | PU-bring:-PT-3s | hospital |
| 'Then I went home.' | | | 'He took me to the hospital.' | | |
- (341) Sāravi (SY: 54)
- | | | | | | | |
|-----------------------------------|-----------------|-----------|---------------------------------|------------|----------------|---------------|
| <i>mimun-a</i> | <i>burd-ene</i> | <i>še</i> | <i>sere.</i> | <i>men</i> | <i>b-ur-em</i> | <i>bazar.</i> |
| guest-PL | go:PU:PT-3P | own | house | I | PU-go-1s | market |
| 'The guests went to their house.' | | | '...(that) I go to the bazaar.' | | | |

Goals in Lazard's corpus are 84.4% postverbal. Goals in Miller's Leriki corpus are 87.9% postverbal, predominantly *bæ*-marked. The following Leriki examples are from the texts gathered in my own field work in 2002–2006:

- (342) Leriki
- | | | | | |
|--|------------------|-------------------------|-------------|--------------|
| <i>še-m</i> | <i>bæ</i> | <i>sivers</i> | <i>bæ</i> | <i>væyæ.</i> |
| go:PT-1s ₁ | to | P.N. | to | wedding |
| 'I went to <i>Si-Vers</i> to a wedding.' | | | | |
| <i>telli</i> | <i>toki=aedæ</i> | <i>omaæ-Ø</i> | <i>ijo?</i> | |
| P.N. | dark=in | come:PT-3s ₁ | here | |
| 'Did Telli come here in the dark?' | | | | |

- (343) Māsulei (LM: 56, IV, 19) (LM: 34, I: 7)
 tə Ø-bær-əm čəmən kæ. še-næ če otay-e xab.
 you PR-take-1S₁ my house go:PT-3P₁ his room-EZ sleep
 'I will take you to my house.' 'They went to his bedroom.'
- (344) Vafsi
 pæs búri-a vəws! bǽ-resa-Ø bæ yeý hæsar.
 then come:IMPER-2P₁ P.N. PU-arrive:PT-3S to one yard
 'Then come to Vafsi!' 'He got to a courtyard.'

The following are some frequency counts for tokens of Recipients and Addressees from texts in four languages covered in this chapter:

Language	Recipients		Addressees	
	Preposed	Postposed	Preposed	Postposed
Kordkheyli	19 90.5 %	2 9.5 %	15 100 %	0 0 %
Māsulei	24 100 %	0 0 %	33 100 %	0 0 %
Leriki	19 60.3 %	14 39.7 %	2 66.7 %	1 33.3 %
Vafsi	32 33.7 %	63 66.3 %	51 96.3 %	2 3.7 %

10. Sample texts with interlinear glossing

The following are four short texts as examples of naturalistic speech in Lāhijāni, Sāravi, Leriki, and Vafsi.

10.1. Lāhijāni: a reminiscence, Faramarz Vakof

The Lāhijāni /a/ and /ə/ phonemes are transcribed here with subphonemic pronunciations.

zæmán-i ki mu kəlas-e no-e debiristan issæ=bu-m,,
 time-RESTR SUB I grade-EZ nine-EZ high_school BE₃=AUX:PT-1S
 muəllim né-ysæ=bu-Ø,
 teacher NEG-BE₃=AUX:PT-3S

When I was in ninth grade of high school, the teacher wasn't there,

n-áémæ=bu-Ø. libas-ə æmra hæminjuri bū-šo-im
 NEG-come:PT=AUX:PT-3S clothes-REZ with just_like_this PU-go:PT-1P
 yek-tə zæmín-ə dirin ki
 one-UNC land-REZ into SUB

he hadn't come. We went just like that with (our regular) clothes on into a field that

hævali-ø *mædrese* *bu-Ø.* *bú-šo-im*
 proximity-EZ school be:PT-3S PU-go:PT-1P
 was near the school. We went
futbal *bazí* *bù-kon-im.* *bæ:d* *æz* *nim* *saæt*
 soccer game PU-do-1P after from half hour
 to play football. After half an hour,

ye *dæfæri* *mu* *bü-ket-im=o*
 one time I PU-fall:PT-1S=and
 all of a sudden I fell down and
mi *zanu-e* *čep* *xéyli* *derd* *kùn-i.*
 my knee-EZ left very pain do-3S
 my left knee was hurting badly.

nigá *b-ùd-əm* *bá-de-m* *yε* *tike* *pare* *sæng* *bá-šo-Ø*
 look PU-do:PT-1S PU-see:PT-1S one piece piece stone PU-go:PT-3S
mi *zanu* *mien.*

I looked (and) saw (that) a little piece of stone pierced my knee.

u *zæmon* *tilifon=o* *i* *čiz-on* *né-dašt-im.* *bæ:d*
 that time telephone=and this thing-PL NEG-have:PT-1P then
yek=tæ *æz* *zak-on-a*
 one-UNC from child-PL-RE

In those days we didn't have telephones and these things. Then I got on the

taerk-e *čærx* *bá-ništ-əm* *æ* *m-æ* *bá-värd-i* *bimarestan.*
 back-EZ bike PU-sit:PT-1S he I-RA PU-bring:-PT-3S hospital
duktør *niga* *b-ud-i*
 doctor look PU-do:PT-3S
 back of one of the guys' bikes (and) he took me to the hospital. The doctor looked

bə-di-Ø *xéyli* *væxim=I.*
 PU-see:PT-3S very serious-COP:3S
 and saw that it's very serious.

bæ:d *æz* *du* *saet* *æmæl-e* *jærahi* *æmæ* *bá-d-im*
 after from two hour operation-EZ surgical we PU-see:PT-1P
 After two hours of surgery, we saw that

ki *mi* *pier=əm* *b-áémæ-Ø.* *m-æ* *sævar-i* *mašin*
 SUB my father=also PU-come:PT-3S I-RA mounted-EZ car
b-úd-i *bá-värd-i* *xoni.*
 PU-do:PT-3S PU-bring:PT-3S house
 our (i. e., my) father also came in. He put me in the car (and) brought (me) home.

mæjbur=a bo-m čar mah æsa-yə æmra rá bù-šo-m,
 forced=CS be:PT-1s four month cane-REZ with road PU-go:PT-1s
 I was forced to walk with a cane for four months.

bæ:tha mutəvæje bu-bo-m ki
 later attentive PU-be:PT-1s SUB
 Later on I realized that

ægæ bá-xa-m futbol bazi bù-kun-əm, baási xéyli ihtiát
 if PU-want-1s soccer game PU-do-1s must:PT very caution
 if I want to play football, I had to be very careful.

bù-kun-əm. hæmin ihtiat
 PU-do-1s this.very caution
 Just this caution

baés bù-bo-Ø mu de ná-ton-əm futbal bazi bù-kun-əm.
 cause PU-be:PT-3s I more NEG-can-1s soccer game PU-do-1s
 caused me not to be able to play football any more.

čon hæmiše i ters
 since always this fear
 Since there was always a fear

ye júr-i hænnæ=bu-Ø ki áge mu dubare bá-ke-m.,
 one way-INDEF BE_s=AUX:PT-3S SUB if I again PU-fall-1s
momken=i bazhæm
 possible=COP:3S again
 somehow that if I should fall again, perhaps

i jærahæt-i ya zæxmí ye číz-i bú-bu-Ø. in
 one surgery-INDEF or injury one thing-INDEF PU-be:SBJ-3S this
bais bú-bu-Ø ki be-ja=ye
 cause PU-be:PT-3S SUB to-place-EZ
 there would be a surgery or an injury (or) something again. This caused me

in-ki mu herekæt bišter bú-kun-əm., bišter bú-š-əm
 this-SUB I movement more PU-do-1s more PU-go-1s
kitabxani=o i jur čiz-on.
 library=and this kind thing-PL
 to go more to the library and things like this instead of moving around more.

10.2. Sāravi (from Yoshie 1996: 64)

zena ruz-a pæmbe golek-a=re Ø-get-e Ø-ši-ye kele
 wife day-PL cotton wad-PL=RA IMF-say:PT-3S IMF-go:IMF-3S river
lu, un-a=re Ø-řesi-ye.
 lip it-PL=RA IMF-spin:PT-3S

Every day the wife would take the cotton wads, go the river edge (and) spin them.

væk-a kele lu qorqor Ø-kard-ene, ve=re xial=e
 frog-PL river lip croak IMF-do:PT-3P s/he=RA imagination=COP:3S
 ke
 SUB

The frogs were croaking on the river bank (and) she imagined [exp. vb.] that

ven-e kar=je irad ger-n-ene. ve=re læj bæ-it-e
 s/he-REZ work-from criticism take-PR-3P s/he=ra spite PU-take:PT-3S
 pæmbe-ha=re dæ-kard-e
 cotton-PL=RA PVB-do:PT-3S

they were (lit: are) criticizing her work. She got spiteful [exp. vb.] (and) dumped the cotton

kele-e u dele. bá-ut-e báe-ir-in, áge šema beter
 river-REZ water inside PU-say:PT-3S PU-take-2P if you better
 bæled=eni báe-res-in,
 able=COP:2P PU-spin-2P

(wads) into the river water. She said, ‘Take them, if you can spin (them) better,

báe-res-in! væk-a æz tærs báe-pperess-ene u-e dele.
 PU-spin-2P frog-PL from fear PU-jump:PT-3P water-REZ inside
 pæmbe golek-a=m
 cotton wad-PL=too

you spin (them)! Out of fear, the frogs jumped into the water. And the cotton wads

b-urd-e u-e tæ. b-urd-e æ-tta ču
 PU-go:PT-3S water-REZ bottom PU-go:PT-3S one-UNC stick
 bæ-it-e u=re behæm Ø-zu-Ø
 PU-take:PT-3S water-RA together IMF-hit:PT-3S

went to the bottom of the water. She went, got a stick (and) was stirring up the water

ſe pæmbe=re bæ-ir-e. væk-a=re Ø-gat-e men ſeme
 own cotton-RA PU-take-3s frog-PL=RA IMF-say:PT-3S I you:PL
 sere=re xerab ka-m-be
 house=RA ruined do-PR-1s

to get her cotton (back). She kept saying to the frogs, ‘I’ll ruin your house [a standard curse]

me pæmbe-ha=re há-d-in!
 my cotton-PL=RA PVB-give-2P
 Give me my cotton!

hæminti=ke u=re læt Ø-zu-Ø, ven-e mærdi
 just.as=SUB water=RA thrash IMF-hit:PT-3S s/he-REZ husband
 sær bæ-resi-ye. bæ-di-ye
 NVE PU-arrive:PT-3S PU-see:PT-3S

‘Just as she was thrashing the water, her husband suddenly arrived. He saw

æ-tta či u-e dele bærq-bærq zæn-d-e.
 one-UNC what water-REZ inside shine-shine hit-PR-3S
 something in the water was (lit: is) shining.

b-úrd-e bæ-it-e, bæ-di-ye
 PU-go:PT-3S PU-take:PT-3S PU-see:PT-3S
 He went, picked (it) up, (and) saw

xešt-e tela=ye. čæn-ta=re ſe bæ-it-e čæn-ta=re
 brick-EZ gold=COP:3S few-UNC=RA self PU-take:PT-3S few-UNC=RA
 há-da-e zena-e sær=o
 PVB-give:PT-3S wife-REZ head=and

they were (lit: are) gold bricks. He picked up a few of them himself and a few of them he put on his wife’s head and

b-úrd-ene sere. mærdi ruz-a xešt-a=re Ø-værd-e
 PU-go:PT-3P house man day-PL brick-PL=RA IMF-take.to:PT-3S
 bazar pul Ø-kard-e,

bazaar money IMF-do:PT-3S
 they went home. Every day the man would take the bricks to the bazaar, changed (them) into money,

Ø-yard-e sendux-e dele ja Ø-da-e. zena=re
 IMF-bring:PT-3S chest-REZ inside NVE IMF-give:PT-3S wife-RA
 Ø=gat-e dæs ná-zen.
 IMF-say:PT-3S hand NEG-hit
 he would bring (it home and) hide (it) in a chest. He would say to his wife, Don’t touch (it)!

10.3. Lerik text: encounters with a Bear, Xanbala Musayev, Pirasora village

mī iimsor avgust-æ mang=ædæ řæ=b-im bæ višæ.
 I:OB this.year August-LINK month-LOC go:PT=AUX:PT-1S₁ to forest
 This year in the month of August I had gone to the forest.

dum=o mī glæ čk-æ spæ híst=e.
 back=ABL I:OB UNC good-LINK dog BE₂=AUX:3S₁
 Behind me is a good dog.

višæ dilæ=dæ i kæræ=n bæ-n-é=m
 forest inside=LOC one time=too FUT-know-INF=1S₁
 In the forest all of a sudden I see
glæ-y ři be-še-Ø.
 UNC-INDF thing out-go:PT-3S₁
 something came out.

dié=m kárde bæ-n-é=m hírs=e. ha næisæ
 look=1S₂ do:PT FUT-know-INF=1S₁ bear=3S₁ yes something
 I looked and I see it is a bear. Well, really,

č-ím-i nav=o vit-e æ-zí-n-im.
 ABL-DEM:PROX-OB front=ABL run-INF DUR-know-NEG-1S₁
 I won't be able to outrun it.

bæ do békü-m, ím=aen
 to tree PVB-go-1S₁ DEM:PROX-too
 If I climb a tree, it also

be-bæ-š-e-Ø. bæ-n-é=m hežo om-é-dæ-Ø.
 PVB-FUT-go-INF-3S₁ FUT-know-INF=1S₁ always come-INF-LOC-1S₁
 will climb up. I see it is still coming.

dí dæhšæti
 with fright
 Frighteningly

p-i-dæ=ž=e bí-kšt-o-Ø odæm-i.
 want-INF-LOC=3S₂=AUX SBJ-kill-OPT-3S₁ person-OB
 it was going to kill you (lit: a person).

sípæ sáe-gní-ye č-æy.
 dog PVB-fall:PT-3S₁ ABL-DEM:DIST-OB
 Well, the dog went after it.

sipæ=n kišt-é-dæ=be. *sipæ=n dàst ni:kæšæ-y.*
 dog=too kill-INF-LOC=AUX:PT-3S₁ dog=too hand NEG-pull:PT-3S₁
 It was going to kill the dog, too. And the dog did not give up.

vít-e.

flee:PT-3S₁
 It (the bear) ran away.

peši č-æ-y se čo ruž-i bæ digæ, še-m
 then ABL-DEM:PROX-OB three four day-OB to other go:PT-1S₁
ižæn rást omæ-y.

again meeting come:PT-3S₁

Then three or four days after that, I went (and) it encountered (me) again.

ovči b-é-bæ-y-Ø, tifang=im pe-gætæ-bæ-y, az=aen
 hunter SBJ-DUR-be:PT-IRR-3S₁ rifle=1S₂ PVB-take:PT-be:PT-IRR I=too
gásti

maybe

If there had been a hunter, (or) if I had picked up a gun, I maybe

b-é-zn-i-m žæ-y. ænjæx œv... čæy gužd
 SBJ-DUR-know-IRR-1S₁ hit:PT-INF only it DEM:DIST:GEN meat
čæy pi
 DEM:DIST:GEN fat

could have shot it, too. Only it... it seems they say its meat (and) its fat

güya vót-dæ=n dæmon=e. ovči omæ-y
 supposedly say-LOC=3P₁ medicine-COP:3S₁ hunter come:PT-3S₁
vot=e isæ bít-žæn-o-m.,
 say-AUX now SBJ-hit-OPT-1S₁
 is supposedly medicinal. A hunter came along (and) said, if I shoot it now,

vot=e gám=e. čæy gužd bú bæ-kà-Ø.
 say-AUX warm-COP:3S₁ DEM:DIST:GEN meat smell FUT-make:PT-3S₁
 he said, it is hot. The meat will go bad (i. e., will start to smell).

10.4. Vafsi text

buælisina yey řu ræfiq bæ-v-e yek-i henra æz iran
 Avicenna one day friend PU-be-:PT-3S₁ one-OB with from Iran
æ-čí-nde
 DUR-go:PT-3P₁
 Avicenna once made friends with someone (as) they were going from Iran

mæmlekæte xareji. b-ísdi yey næfær šaér-i yey
 country-EZ foreign PU-3S₂-see:PT one HCL poet-INDEF one
ketáb-i vær=es=dær=e.
 book-INDEF side=3S₂=LOC=COP:3S₁

to a foreign country. (There) he saw a poet (and) there was a book next to him.

bawæli in kætab hágiræ hey dær=es-gærd-en-a æz
 Avicenna this book PVB-took-PT REPET PVB=3S₂-turn-CAUS-PT from
ævvæle ketáb-i
 first-EZ book-OB

Avicenna picked up the book and kept leafing through it from beginning

yawa axære ketáb-i murur=es a-kærd=o hár=esda
 until last-EZ book-OB going.over=3S₂ PVB-do:PT=and PVB=3S₂-give:PT
o saab ketáb-i.
 to owner book-OB

to end of the book, and then he gave it back to its owner.

héni-ke bá-ssi-nde ša-y palu in šaere-y kætab tæqdim
 when-SUB PU-go:PT-3P₁ shah-OB side this poet-OM book present
kærd-e ša-y.
 do:PT-PERF? shah-OB

When they went to the Shah, this poet presented the book to the king.

ša-y hágiræ naw=es bá-kærd=o va-Ø,
 shah-OB PVB-take:PT look=3S₂ PU-do:PT=and say:INTR:PT-3S₁
 The Shah took it and looked at it and said,

xéyli næč kætab=e.
 very good book=COP:3S₁
 'This is a really nice book.'

bawælisina va, aqa, an kætab tæmen=e,
 Avicenna say:INTR:PT-3S₁ sir that book I:OB=COP:3S₁
tini ne.
 he-OB COP:NEG:3S₁
 'Avicenna (then) said, sire, that book is mine, it is not his.'

in šaeræ va-Ø ke an duru r-vaz-e.
 this poet say:INTR:PT-3S₁ SUB DEM:PRON:DIST lie DUR-say-3S₁
 The poet (then) said (that), 'He is lying!'

šeš sal há-nešeste-ym
 six year PVB-sit:PT-1S₁
 For six years I sat (and)

zæ:mæt=om bæ-keša kætab=om bæ-nnevisde.
 trouble-1s₂ PU-pull:PT book=1s₂ PU-write:PT
 labored (and) I wrote the book.

tæ æte r-vaz-i tæmen=e?
 you now DUR-say-2s₁ I:OB=COP:3s₁
 Now you say ‘it’s mine’?

*va-Ø, æz hær sætr tæ há-pærs-i ke či
 say:INTR:PT-3s₁ from every line you PVB-ask-2s₁ SUB what
 nevisdæ, æz pak æz bær*

write:PT:PPL I all from memory

He said, ‘Any line you ask me (about) what is written there, I will tell it all

*ær-vaz-om veréwn=i=ra.
 DUR-say-1s₁ OPBII=2s₂=for
 to you by heart.’*

*æz hær væræq=esan dæ-gerd-en-a,, buælisina-y tæman
 from every sheet=3p₂ PVB-turn-CAUS-PT Avicenna-OM all
 Whatever page they turned to, Avicenna*

*æz bær
 from memory
 recited it all by heart.*

*bæ-xand. šaer-æ va-Ø ke tæmen ba-watt-e æmma
 PU-read:PT poet-DEF say:PT-3s₁ SUB I:OB PU-say:PT-PERF but
 The poet said, ‘I have written (lit: said) it but*

*vir=om=dæ næ-mænd-e. ša va-Ø ke kætab pæs
 memory=1s₂=in NEG-stay:PT-PERF king say:PT-3s₁ SUB book then
 tini=e esdæ n-e.
 he:OB=COP:3s₁ you:OB COP:NEG:3s₁*

I don’t remember it (by heart).’ The Shah said, ‘then the book is his, not yours.’

*bæzáni buælisina va-Ø, næ, kætab tini=e.
 then Avicenna say:PT-3s₁ no book he:OB=COP:3s₁
 Then Avicenna said, ‘no, the book is his.*

*tæmen man kæšdi=dæ hár=om-giræ
 I:OB inside boat=in pvb=1s2-take:pt
 I picked it up on the boat*

*æz bala yawa pain murur=om kærð-e,
from top until bottom review=1s₂ do:PT-PERF
hefz=om kærð-e.
memorization=1s₂ do:PT-PERF
in the boat (and) leafed through it from top to bottom and memorized it.'*

Abbreviations

1P	1pl. PAM	CS	Change of State
1P ₁	1pl. PAM, Set ₁	F	feminine
1P ₂	1pl. PAM, Set ₂	DEF	Definite
1S	1sg. PAM	DEM	Demonstrative
1S ₁	1sg. PAM, Set ₁	DF	Direct Feminine
1S ₂	1sg. PAM, Set ₂	DF ₂	Direct Feminine II
2P	2pl. PAM	DIR	Direct Case
2P ₁	2pl. PAM, Set ₁	DIST	Distal
2S	2sg. PAM	DP	Direct Plural Case
2S ₁	2sg. PAM, Set ₁	DUR	Durative
2S ₂	2sg. PAM, Set ₂	EMF	Emphatic
3P	3pl. PAM	EMOT	Emotive particle
3P ₁	3pl. PAM, Set ₁	EMPF	Emphatic
3P ₂	3pl. PAM, Set ₂	EZ	<i>Ezafe</i> linker
3S	3sg. PAM	FEM	Feminine
3S ₁	3sg. PAM, Set ₁	FUT	Future
3S ₂	3sg. PAM, Set ₂	GEN	Genitive
3SF	3sg. Feminine	IMF	Imperfect
3SF ₁	3sg. Feminine Set ₁	IMPER	Imperative
3SM	3sg. Masculine	INDF	Indefinite
3SM ₁	3sg. Masculine Set ₁	INF	Infinitive
ABL	Ablative	INTR	Intransitive
ACC	Accusative	IRR	Irrealis
ACC ₁	Accusative, type 1	JE	Polysemous adposition
ACC ₂	Accusative, type 2	KA	Polysemous adposition
ÆDÆ	polysemous adposition	KU	Polysemous adposition
ALL	Allative	LINK	General NP linker
AUX	Copula as Auxiliary	LOC	Locative
BE ₁	'Be' verb One	LV	Light Verb
BE ₂	'Be' verb Two	M	Masculine
BE ₃	'Be' verb Three	NEG	Negative
BE ₄	'Be' verb Four	NVE	Non-verbal element in light verb construction
BE ₅	'Be' verb Five	OB	Oblique
BE _{DEM}	'Be' based on Demon- strative formant	OBJ	Objective Case
CAUS	Causative	OF	Feminine Oblique
COND	Conditional	OF ₂	Feminine Oblique II
COP	Copula	OM	Masculine Oblique

OP	Plural Oblique	PROX	Proximal
OPBII	Oblique Pronominal	PT	Past
	Base II (Benefact.)	PU	Punctual
OPT	Optative	PVB	Preverb
PL	Plural	RA	RA Objective marker
OS	Singular Oblique	REPET	Repetitive particle
PAM	Person-Agreement Marker	RESTR	Restrictive marker
PAM ₁	PAM Set ₁		for Relative clauses
PAM ₂	PAM Set ₂	REZ	Reverse <i>Ezafe</i>
PASS	Passive	SBJ	Subjunctive
PD	Direct Plural	SD	Singular Direct
PERF	Perfect	SDF	Sg Direct Feminine
PL	Plural	SET ₁	Set of Direct PAMs
PO	Oblique Plural	SET ₂	Set of Oblique PAMs
POSS	Possessive	SG	Singular
PPL	Past Participle	SOF	Sg Oblique Feminine
PR	Present	SOM	Sg Oblique Masculine
PR ₁	Present I	STAT	Stative/Resultative
PR ₂	Present II	SUB	Universal Subordinator
PROG	Progressive	TEMP	Temporal
PROH	Prohibitive	UNC	Universal Numeral
PRON	Pronoun		Classifier

References

- (AF=) 'Abdoli, 'Ali. 2001/1380. *Farhang-e tatbiqi-e tâleši, tâti, āzari* [translation missing]. Tehran: šerkat-e sahâmi-e entešâr.
- 'Abdoli Koluri, 'Ali. 1992/1371. *Nazari be jâme'e-ye 'ašâyeri-e tâleš* [A glimpse into the tribal community of the Talesh]. Tehran: Etelâ'āt Publishers.
- (AA2=) Al-e Ahmad, Jalal. 1958/1337. *Tânešinhâ-ye boluk-e Zahrâ* [The Tat population of Zahrâ District]. Dâneš: Tehran.
- (AA1=) Al-e Ahmad, Jalal. 1973/1352. *Owrâzân*, 4th edn. Mâziār: Tehran.
- Amini, Tanja J. 2003. Das Tonekâbonî von Šîrûd: Materialien und Untersuchungen zur Annäherung an einen kaspischen Dialekt. Hamburg: University of Hamburg, unpublished MA thesis.
- Amirzâdeh-Vâjârgâh, Asghar. 1996/1375. *Ārus dâmât sêng-e siyâ bobon* (*Arus va dâmâd sang-e siâh şodand*) [The bride and groom turned into a black stone]. Rašt: Našr-e Gilakân.
- Arkadiev, Peter M. 2008. Poor (Two-term) case systems: Limits of neutralization. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook on case*, 686–99. Oxford: Oxford University Press.
- Asatrian, Garnik & Habib Borjian. 2005. Talish and the Talishis (The state of research). *Iran and the Caucasus* 9(1). 43–72.
- Bazin, Marcel. 1981. Quelques échantillons de variations dialectales du Tâleši. *Studia Iranica* 10. 111–124, 269–277.

- Borjian, Habib & Maryam Borjian. 2007. Marriage rites in South Caspian villages. *Archiv orientální*, vol. 75(2), 191–214.
- Borjian, Maryam & Habib Borjian. 2007. Ethno-linguistic materials from rural Mazandaran: Mysterious memories of a woman. *Iran and the Caucasus* 11(2). 226–254.
- Borjian, Habib & Maryam Borjian. 2008. The last Galesh herdsman: Ethno-linguistic materials from South Caspian Rainforests. *Iranian Studies* 41(3). 365–402.
- Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77. 433–455.
- Grunberg, Aleksandr L. 1963. *Jazyk Severoazerbajdžanskix Tatov* [The language of the northern Azerbaijani Tat]. Moscow: Izdatel'stvo Akademii Nauk SSSR.
- Guizzo, Daniele. 2004. *Due dialetti Taleshi di Talesh (Hashtpar)*. Naples: University of Naples dissertation.
- Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A construction grammar approach*. Berlin & Boston: De Gruyter Mouton.
- (HA=) Hāsemi, Jahāngir. 1974/1353. *Guyeš-e Alviri* [The Alviri dialect]. Tehran: University of Tehran dissertation.
- Ivanow, Wladimir. 1931. The dialect of Gozärkhon in Alamut. *Acta Orientalia* 9. 352–369.
- Jahāngiri, 'Ali Asghar. 1988/1367. *Kandelus*. Tehran: Mo'assese-ye farhangi-e Jahāngiri.
- (NJ=) Jahāngiri, Nāder. 2003. *Guyeš-e Gilaki-e Lāhijān, Vol I, vižegihā-ye āvāyi va sāxt-vāžei* [The Gilaki dialect of Lāhijān, Vol. I. The phonological and morphological characteristics]. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Kalbāsi, Irān. 1997/1376. *Guyeš-e Kelārdašt (Rudbārak)* [The dialect of Kelārdašt (Rudbārak)]. Tehran: Pažuheşgāh-e 'olum-e ensāni va motale'āt-e farhangi.
- (AAK=) Kārang, 'Abdol-'Ali. 1954/1333. *Tāti va Harzani*. Tabriz: Esma'il Va'ezpur.
- Kaye, Stephen. 2013. Morphemic stems in the northern Talyshi verb: Diachrony and synchrony. In Silvio Cruschina, Martin Maiden & John Charles Smith (eds.), *The boundaries of pure morphology: Diachronic and synchronic perspectives*, 181–208. Oxford: Oxford University Press.
- Lambton, Anne K. S. 1938. *Three Persian dialects*. James G. Forlong Fund, 18, London: Royal Asiatic Society.
- (LM=) Lazard, Gilbert. 1979. Textes en tâlesi de Mâsule. *Studia Iranica* 8(2). 33–66.
- Lazard, Gilbert. 1990. Le dialecte de Rudbâr (Gilân). In Dina Amin & Manouchehr Kasheff (eds.), *Iranica Varia: Papers in honor of Professor Ehsan Yarshater*, 110–123. (Acta Iranica 30. Troisième Série, vol. XVI. Textes et Mémoires). Leiden & Boston: Brill.
- Lecoq, Pierre. 1989. Les dialectes du centre de l'Iran. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 313–326. Wiesbaden: Reichert.
- Farzpur-Māčiāni, M. 1964, 1965. Zabān-o farhang-e Māčiān [The language and lexicon of the Mačiān], *Našriyya-ye Daneškada-ye Adabiyyāt-e Tabriz* 16, 1964/1343: 277–296, 451–470; 17, 1965/1344: 109–128, 261–284.
- (MT2=) Miller, Boris Vsevolodovič. 1930. *Talyšskie teksty* [Talyshi texts]. Moscow: Naučno-issledovat'skij institut nacional'nyx i e'tničeskix kul'tur narodov vostoka SSSR.
- Miller, Boris Vsevolodovič. 1953. *Talyšskij Jazyk* [The Talyshi language]. Akademiya Nauk: Moscow.
- (MH=) Mortazavi, Manuchehr. 1954/1333. Noktei cand az zabān-e harzani [Some points on the Harzani language]. *NDATabriz* 6. 304–314.

- (MH=) Mortazavi, Manuchehr. 1962/1341. *Fe'l dar zabān-e harzani* [The verb in the Harzani language]. *NDATabriz* 14. 453–488.
- (MH=) Mortazavi, Manuchehr. 1963/1342. *Fe'l dar zabān-e harzani* [The verb in the Harzani language]. *NDATabriz* 15. 61–97.
- Nawata, Tatsuo. 1982. The Masal dialect of Talishi. *Acta Iranica* 22. 93–117.
- Pahlavān, Keyvān. 2003/1382. *Farhang-e mardom-e Elāšt (zādgāh-e Rezā šāh) va Savādkuh* [The culture of the people of Elāšt (the birthplace of Rezā Shāh) and Savādkuh]. Tehran: Ārvan/Xāvarzamin.
- Paul, Daniel. 2011. *A comparative dialectal description of Iranian Taleshi*. Manchester: University of Manchester dissertation.
- Pāyandeh, Mahmud. 1973/1352. *Masalhā va estelāhāt-e Gil va Deylam* [Dictionary of proverbs and expressions of Gilan and Deylam]. Tehran: Bonyād-e farhang-e Iran.
- Pāyandeh, Mahmud. 1987/1366. *Farhang-e Gil va Deylam, Fārsi be Gilaki* [Dictionary of Gil and Deylam, Farsi to Gilaki]. Tehran: Amirkabir.
- Pirejko, L. A. 1976. *Talyšsko-russkij i russko-talyšskij slovar* [Talyshi-Russian, Russian-Talyshi dictionary]. Moskva: Izdatel'stvo Russkij jazyk.
- (RR=) Rastorgueva, Vera S., Aza A. Kerimova, Ahmed K. Mamedzade, L. A. Pirejko & D. I. Èdel'man. 1971. *Giljanskij jazyk* [The Gilaki language]. Moscow: Nauka.
- Rastorgueva, Vera S. & D. I. Èdel'man. 1982. Giljanskij, mazanderanskij (s dialektami šamerzadi i velatru) [Gilaki, Mazanderani (including the dialects of Shahmirzad and Velatru)]. In Vera S. Rastorgueva (ed.), *Osnovy iranskogo jazykoznanija, Novoiranskie jazyki: Zapadnaja gruppa, Prikaspisjkie jazyki*, 447–554. Moscow: Nauka.
- Redard, Georges. 1970. Other Iranian languages. In Thomas Sebeok (ed.), *Current trends in linguistics, vol. 6, Linguistics in South West Asia and North Africa*, 97–135. The Hague & Paris: Mouton.
- (SH=) Shokri, Guiti. 1995/1374. *Guyeš-e Sāri (Māzanderāni)* [The dialect of Sāri (Māzanderāni)]. Tehran: Pažuhešgāh-e 'olum-e ensāni va motale'āt-e farhangi.
- (SR=) Shokri, Guiti. 2006/1385. *Guyeš-e Rāmsari* [The Rāsari dialect]. Tehran: Pažuhešgāh-e 'olum-e ensāni va motale'āt-e farhangi.
- Sokolova, Vera. 1953. *Očerki po fonetike iranskix jazykov* [Sketches of Iranian Phonetics]. Moscow: Izdatel'stvo Akademii Nauk.
- Stassen, Leon. 2005. Predicative possession. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press.
- Stilo, Donald L. 1981. The Tati language group in the sociolinguistic context of northwestern Iran and Transcaucasia. *Iranian Studies* 14. 137–187.
- Stilo, Donald L. 1994. Phonological systems in contact in Iran and Transcaucasia. In Mehdi Marashi (ed.), *Persian studies in North America: studies in honor of Mohammad Ali Jazayery*, 75–94. Bethesda, Maryland: Iranbooks.
- Stilo, Donald L. 2001. Gilan, x. Languages. In Ehsan Yarshater (ed.), *Encyclopædia Iranica* 10.6, 660–668. New York: Center for Iranian Studies, Columbia University.
- (VF=) Stilo, Donald L. 2004a. *Vafsi folk tales*. Wiesbaden: Reichert.
- Stilo, Donald L. 2004b. Coordination in three Western Iranian languages: Vafsi, Persian and Gilaki. In Martin Haspelmath (ed.), *Coordination*, 269–330. Amsterdam & Philadelphia: Benjamins.
- Stilo, Donald L. 2005. Iranian as buffer zone between the universal typologies of Turkic and Semitic. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic con-*

- vergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 35–63. London & New York: Routledge Curzon.
- Stilo, Donald L. 2007a. Isfahan, xxi. provincial dialects. In E. Yarshater (ed.), *Encyclopaedia Iranica* 14.1, 93–112. New York: Center for Iranian Studies, Columbia University.
- Stilo, Donald L. 2007b. Isfahan, xix. Jewish dialect. In E. Yarshater (ed.), *Encyclopaedia Iranica* 14.1, 77–84. New York: Center for Iranian Studies, Columbia University.
- Stilo, Donald L. 2007c. Isfahan, xxii. Gazi dialect. In E. Yarshater (ed.), *Encyclopaedia Iranica* 14.1, 112–119. New York: Center for Iranian Studies, Columbia University.
- Stilo, Donald L. 2008a. Two sets of mobile verbal person agreement markers in the Northern Talyshi language. In Simin Karimi, Vida Samiian & Donald Stilo (eds.), *Aspects of Iranian linguistics: Papers in honor of Mohammad Reza Batani*, 363–390. Newcastle-upon-Tyne: Cambridge Scholars Press.
- Stilo, Donald L. 2008b. Case in Iranian: From reduction and loss to innovation and renewal. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook on case*, 363–390. Oxford: Oxford University Press.
- Stilo, Donald L. 2009. Circumpositions as an areal response: The case study of the Iranian zone. In Lars Johanson (ed.), *Turkic Languages*, 3–33. Wiesbaden: Harrassowitz.
- Stilo, Donald L. 2010. Ditransitive constructions in Vafsi: A corpus-based study. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook*, 243–76. Berlin & New York: De Gruyter Mouton.
- Stilo, Donald L. 2012. Intersection zones, overlapping isoglosses, and ‘fade-out/fade-in’ phenomena in central Iran. In Behrad Aghaei & Mohammad R. Ghanoonparvar (eds.), *Iranian languages and culture: Essays in honor of Gernot Ludwig Windfuhr*, 3–33. Costa Mesa: Mazda.
- Stilo, Donald L. 2014. Further Notes on the Iranian Substratum of Azerbaijani Turkish. In Heidi Stein (ed.), *Turkic Language in Iran – Past and Present*, 271–293. Wiesbaden: Harrassowitz.
- Stilo, Donald L. 2015. The polygenetic origins of the Talyshi language. In Uwe Bläsing, Victoria Arakelova & Matthias Weinreich (eds.), *Studies on Iran and the Caucasus: In honour of Garnik Asatrian*, 411–453. Leiden & Boston: Brill.
- Stilo, Donald L. 2016. On the Non-Persian Iranian Substratum of Azerbaijan. In Éva Á. Csató, Lars Johanson, András Róna-Tas & Bo Utas (eds.), *Turks and Iranians: Interactions in Language and History*, 185–234. Wiesbaden: Harrassowitz.
- Stilo, Donald L. 2018a. Numeral classifier systems in the Araxes Iran linguistic area. In William B. McGregor & Søren Wichmann (eds.), *The diachrony of classification systems. Current Issues in Linguistic Theory*. Amsterdam: John Benjamins.
- Stilo, Donald L. 2018b. Dikin Marāqeī Tati of Alamut: an undocumented conservative Tati language. In Salome Gholami (ed.), *Endangered Iranian languages*.
- Stilo, Donald L. forthcoming a. Mazanderani. *Encyclopaedia Iranica*. New York: Center for Iranian Studies, Columbia University.
- Stilo, Donald L. forthcoming b. Loss vs. expansion of gender in Tatic languages: Kafteji (Kabatei) and Kelasi.
- Stilo, Donald L. to appear. *A grammar of Vafsi*.
- Tāhbāz, Sirus. 1963/1342. Yuš. Dāneškadaye Adabiyāt, Entešārātē Mo'assesaye Moṭāle'āt va Tahqiqātē Ejtemā'i, No. 9, Daftarhāye Monogrāfi 2, University of Tehran.
- Windfuhr, Gernot L. 1987. Convergence: Iranian Talyshi and Turkic Azari. In Elena Bashir,

- Madhav M. Deshpande & Peter E. Hook (eds.), *Selected papers from SALA-7. South Asian Roundtable Conference, Bloomington, Indiana*, 385–405.
- Windfuhr, Gernot L. 1989. New West Iranian. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 251–62. Wiesbaden: Reichert.
- Windfuhr, Gernot L. 1992. Cases. In E. Yarshater (ed.), *Encyclopaedia Iranica*, 5.1, 25–37. New York: Center for Iranian Studies, Columbia University.
- Yarshater, Ehsan. 1959. The Dialect of Shāhrud (Khalkhāl). *BSOAS* 22. 52–68.
- (KJ=) Yarshater, Ehsan. 1960. The Tāti Dialect of Kajal. *BSOAS* 23. 275–286.
- (YA=) Yarshater, Ehsan. 1964. The dialects of Alvir and Vidar. In Georges Redard (ed.), *Indo-Iranica: Mélanges présentés à Georg Morgenstierne*, 177–187. Wiesbaden: Harrassowitz.
- (ST=) Yarshater, Ehsan. 1969. *A grammar of southern Tati dialects*. The Hague & Paris: Mouton.
- (YT=) Yarshater, Ehsan. 1970. The Tati dialect of Tārom. In Mary Boyce & Ilya Gershevitch (eds.), *W. B. Henning memorial volume*, 451–467. London: Lund Humphries.
- (ASA=) Yarshater, Ehsan. 1996. The Tāleshi of Asālem. *Studia Iranica* 25(1). 83–113.
- (YX=) Yarshater, Ehsan. 2003. The Xo'ini Dialect. *Persica XIX*, Louvain: Peeters. 165–182.
- (KL=) Yarshater, Ehsan. 2005. Tāti Dialect of Kalāsuri. In Dieter Weber (ed.), *Languages of Iran: Past and present*. *Iranica* 8, 269–284. Wiesbaden: Harrassowitz.
- Yarshater, Ehsan. to appear b. Karingāni Dialect. (February 18, 2003 draft).
- Yarshater, Ehsan. to appear c. Karani Dialect. (May 18, 2003 draft).
- (SY=) Yoshie, Satako. 1996. *The Sāri dialect*. Tokyo: Institute for Languages and Cultures of Asia and Africa.
- Zoka, Yahyā. 1954/1332. *Guyeš-e Keringān* [The dialect of Keringān]. Tehran: Daneš.
- Zoka, Yahyā. 1957/1336. *Guyeš-e Galinqaya ya Harzandi* [The dialect of Galinqaya]. *Faranhang-e Irānzamin* 5, 51–92. Tehran.

In addition to the above sources Dr. Ehsan Yarshater generously provided the author access to his handwritten field notes on Anbarāni, Dizi, Gandomābi, Jowkandāni, Karani, Kalāsuri, Karani, Karnaqi, Khoini, Khoynarudi, Lerdi, Nowkiāni, Qal'ei, Razajerdi, Sayyādlari, Viznei, (Northern Talyshi, Iran), Asālemi, Koluri, Kajali, Gandomābi, Nowkiāni, Karani Gandomābi, Karnaqi, Dizi for the preparation of this article.

Stilo field notes: Āmoli, Asālemi, Bābolsari, Charozh, Dikin-Marāqeи, Dikin-Pashēi, Ezbaram, Gurāni, Jubani, Kafteji, Kelāsi, Khatirābādi, Koluri, Lāhijāni, Langerudi, Lenkoran dialect, Marāqeи, Māsāli, Māsulei, Northern and Southern Talyshi (Stilo: in press), Northern Talyshi, Lerik, Oskulak Talyshi, Tutkāboni/Tutkāboni, Rāmsari, Rashti, Razajerdi, Rostamābādi, Rudbāri, Tonekāboni (unless otherwise marked), Tutkāboni, Urei, Vafsi, Xatirābādi.

6. Caucasian rim and southern Black Sea coast

6.1. Caucasian rim and southern Black Sea coast: overview

Geoffrey Haig and Geoffrey Khan

This section covers three languages: Ossetic (East Iranian, Indo-European), Romeyka (Hellenic, Indo-European), and Laz (Georgian-Zan, Kartvelian). All three are spoken in enclaves, separated from their closest genetic relatives and surrounded by unrelated languages. Geographically, Laz and Romeyka are spoken in the mountainous Black Sea coastal region in today's northeastern Turkey. Laz and Romeyka may have had some mutual influence on each other, and both show evidence of a shared superstrate influence from Turkish. They also share a broadly comparable sociolinguistic context: both are endangered minority languages, spoken by Muslim communities within the Turkish state, with virtually no institutional support or official status. Ossetic, on the other hand, is a recognized regional language spoken in a politically disputed region of the the Caucasus. In recent times, its main contact language is Georgian (Kartvelian), but it bears traces of influence from all three indigenous language families of the Caucasus, as well as from Turkic (Erschler 2009: 422, and this volume, chapter 6.3, §6). It has some status as a regional standard, with books and periodicals in an established written form, and in South Ossetia at least, restricted representation in education as well. However, for all three languages, the general pattern is that very few children are acquiring the languages, so their long term prospects of survival are fragile.

In the areal context of this volume, these three are perhaps best characterized *ex negativo* as being outside of the “Mesopotamian” region (Haig 2017b, Haig and Khan, this volume, chapter 1), to which e. g. the Semitic and Iranian languages of northern Iraq, and eastern Anatolia belong (see this volume, sections 2 and 3). This is perhaps most obviously the case with Ossetic and Laz, which lack many of the features that typify the languages of the Mesopotamian region (for example, both lack prepositions, nominal gender, post-nominal position of adjectives and possessors, and both have at least five case marking suffixes, compared to the minimal presence of suffixal case marking in the Semitic and Iranian languages of Mesopotamia). Romeyka, on the other hand, shares certain features, such as prepositions and nominal gender, with the languages of the Mesopotamian region, though other aspects of its morphosyntax have converged with local varieties of Turkish.

In phonology, Ossetic and Laz both exhibit a signature feature of the Caucasus, namely an additional series of voiceless obstruents, variously referred to as ejective, or glottalized. In Laz, glottalization is a feature inherited from Kartvelian, though

Lacroix (chapter 6.2, §2.1) notes that glottalization in Laz is “not as strong” as in Standard Georgian. Ossetic on the other hand, has developed glottalized obstruents through language contact, rather than inheritance. There are obvious parallels here to the spread of pharyngealization from Semitic into neighbouring Iranian languages; both cases involve what is generally considered a phonetically marked feature, yet which is apparently readily borrowed (see Blevins (2017: 96–109) for discussion of comparable cases in terms of the “perceptual magnet effect”).

Adposition ordering is a feature that has been shown to be sensitive to areal influence (Stilo 2005; Haig 2017), though it changes slowly, i. e. it requires prolonged and intensive contact before an inherited adpositional type will adapt wholesale to a neighbouring, distinct, adpositional type. For the languages to the north and east of our region, we would predict postpositions, and both Laz and Ossetic conform to these expectations. Romeyka on the other hand retains its inherited prepositions.

With regard to alignment in case marking and agreement, no clear overall trend can be discerned. Romeyka is straightforwardly accusative in both its verbal agreement and nominal case marking system, the only complication being the typologically unusual (and not fully consistent) overt marking of masculine singular subjects (regardless of transitivity) through a ‘nominative’ case marker (cf. *o komši-s* ‘the neighbour-NOM’, chapter 6.4, Table 3). Laz on the other hand reflects the well-known complexities of its Kartvelian relatives in case marking and agreement. Subjects may be in the absolute, ergative, or dative case, depending on verbal semantics and transitivity, but also on the tense specification of the verb. While case marking of nominal arguments could be considered to be ergative in alignment, the extensive use of dative subjects with, e. g. experiencers, or with verbs with a potential or abilitative derivation, is not really accounted for within conventional alignment typology. Laz also exhibits the property known as ‘inversion’ in Kartvelian linguistics, whereby the two sets of verbal agreement markers, Set I and Set II, which are respectively controlled by one argument type, swap their controllers, depending e. g. on the tense of the verb. For example, with the present tense verb in (1a), the subject determines agreement by Set I (the suffix *-s*), while the object triggers agreement by Set II (*m-*); in the perfect tense verb in (1b), by contrast, the subject triggers agreement by Set II features (*-u-*), while the object triggers agreement by Set I (*b-*) (example provided by René Lacroix):

- | | |
|--------------------------|---------------------|
| (1) a. <i>m-dzi-om-s</i> | b. <i>b-u-dzi-u</i> |
| 111-see-TH-13SG | 11-113.PERF-see-TH |
| ‘he sees me’ | ‘he has seen me’ |

Superficially, this looks rather similar to the tense-sensitive shift in agreement patterns found in both Northeastern Neo-Aramaic (Khan, this volume, chapter 3.4), and Central Kurdish (Haig, this volume, chapter 3.3), and the similarities have been noted in the literature (Coghill 2016). For example in transitive clauses

in Central Kurdish in the present tense (2a), the subject and object are indexed through a verbal suffix and a pronominal clitic respectively (glossed in (2b) as Set I and Set II), while in the past tense (3b), this is reversed:

- (2) Central Kurdish
- | | |
|----------------------------|----------------------------|
| a. <i>nā-yān-bīn-īn</i> | b. <i>na-mān-dī-n</i> |
| NEG-3PL(II)-see.PRS-1PL(I) | NEG-1PL(II)-see.PST-3PL(I) |
| ‘We don’t see them’ | ‘We didn’t see them’ |

However, while a contact scenario for the Kurdish and Neo-Aramaic similarities appears to be plausible (in view of the close structural parallels, and our knowledge of the histories of the respective speech communities), to what extent contact can be invoked in the Kartvelian developments remains conjectural. Ossetic, on the other hand, has lost all trace the ergativity that is presumed to have characterized middle Iranian (Haig 2017a). The rich case-marking system is entirely accusative, and is not sensitive to verbal tense or aspect. The Ossetic case system has been claimed to reflect “Caucasian” influence, but Erschler (2009) points to the difficulty of identifying the source language. However, Erschler (2009) does identify West Caucasian influence in the development of possessive marking in Ossetic. Likewise, influence from Kartvelian, or from Slavic, or from both, has been suggested for the system of aspectual preverbs in Ossetic (see also chapter 6.3, §6).

Turning to the traditional features of word order typology, we find that all languages have pre-nominal genitives and adjectives, in line with Turkic and the languages of the Caspian region (Stilo, this volume, chapter 5). In this respect, the languages of this section are distinct from all other Iranian languages considered in this volume, and from the Semitic languages. With regard to the order of verb and direct object, both Laz and Ossetic are fairly consistently OV, while the status of Romeyka in this respect remains somewhat unclear. It seems reasonable to assume that Romeyka inherited VO word order, but under Turkish influence may now permit OV, though the conditions for this, and its dialectal distribution within Romeyka itself, remain to be established. Where we do find consistency, however, is in the placement of the copula, which in all three languages is clause final, at least as the unmarked option. Overt, clause-final copulas is a feature that is common to the entirety of East Anatolia (Haig 2017b), and even extends to those Arabic dialects of northern Iraq spoken “east of the Tigris” (Procházka, this volume, chapter 3.2, §2.4.1). Position of the copula is not a feature that is considered in traditional word order typology, but from an areal linguistic perspective, it may be more sensitive to contact influence than, for example, the order of object and verb.

One of the striking features of the Iranian and Semitic languages of the Mesopotamian region is the predominance of finite verb forms in dependent clauses. In the extreme case, languages of Mesopotamia have virtually no non-finite syntax, so that even modals such as ‘be able’, ‘want’, or ‘need’, or same-subject phasal

verbs such as ‘begin’ or ‘finish’, require a finite dependent clauses (see Haig and Khan, this volume, chapter 1). Although the languages of the Caucasian rim do make extensive use of finite forms in dependent clauses, there is a greater tendency towards non-finite verb forms. For example in Ossetic, “the infinitive in *-ən* [...] is used with phasal, modal, emotive, mental, causation, speech and evaluative predicates” (Serdobolskaya 2016: 304). Laz also uses non-finite forms (verbal nouns, chapter 6.2, §4.6) for the verbs ‘forget’ and ‘want’. Similarly, in Romeyka non-finite forms are found with certain types of dependent clauses (chapter 6.4, §6.5.2). Thus the broad generalization is that the strong tendency towards finiteness in dependent clauses of all types is significantly weaker outside of the Mesopotamian region.

Finally, we consider a type of dependent clause known to be sensitive to areal influence, namely relative clauses. The general pattern for all the languages in the other sections of this volume is for post-nominal, finite relative clauses, at least for non-subject relativization. This holds even in the Turkic languages of western Iran and northern Iraq (Bulut, this volume, chapters 3.5 and 4.2), even though standard Turkish has pre-nominal, participial relative clauses. But in the languages of the Caucasian rim, the dominance of post-nominal relative clauses is absent. Ossetic uses a correlative strategy (chapter 6.3, §5.4), which involves a head noun internal to the relative clause, but additionally represented by some pronominal element external to the relative clause. In Laz, the dominant strategy is prenominal (chapter 6.2, §6), and in Romeyka, both prenominal and postnominal relative clauses appear to be possible (chapter 6.4, §6.5.1). While all three languages evidently differ in detail in their relativization strategies, common to all is the possibility of placing the relative clause before the head noun, and in this point, they clearly differ from the languages south of the Caucasian Rim, where post-nominal relativization is preferred, and in several languages represents the sole option.

References

- Blevins, Juliette. 2017. Areal sound patterns: From perceptual magnets to stone soup. In Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 88–121. Cambridge: Cambridge University Press.
- Coghill, Eleanor. 2016. *The Rise and Fall of Ergativity in Aramaic*. Oxford: Oxford University Press.
- Erschler, David. 2009. Possession marking in Ossetic: Arguing for Caucasian influences. *Linguistic Typology* 13(3). 417–450.
- Haig, Geoffrey. 2017a. Deconstructing Iranian ergativity. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *The Oxford Handbook of Ergativity*, 465–500. Oxford: Oxford University Press.
- Haig, Geoffrey. 2017b. Western Asia: East Anatolia as a transition zone. In Raymond

- Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 396–423. Cambridge: Cambridge University Press.
- Serdobolskaya, Natalia. 2016. Semantics of complementation in Ossetic. In Kaspar Boye & Peter Kehayov (eds.), *Complementizer semantics in European languages*, 293–339. Berlin: De Gruyter.
- Stilo, Donald. 2005. Iranian as buffer zone between the universal typologies of Turkic and Semitic. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*, 35–63. London: Routledge.

6.2. Laz

René Lacroix

1. Introduction

Laz belongs to the South Caucasian language family, also known as Kartvelian. The other three members are Mingrelian, Georgian and Svan. The genetic relationships between these languages are displayed in Figure 1. Laz and Mingrelian, which are closely related, were once considered as two varieties of the same language, called Zan.

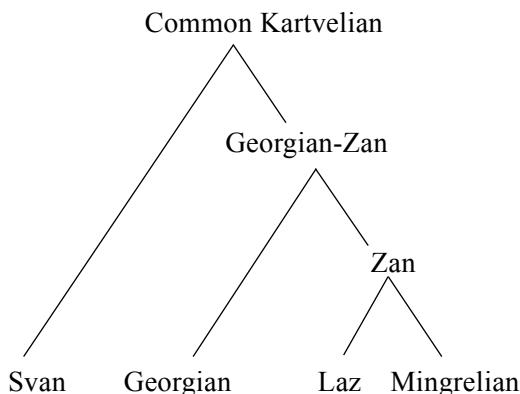


Figure 1: Genetic relationships between the South Caucasian languages

The Laz live predominantly in the mountainous region of northeastern Turkey, on a strip of land along the shore of the Black Sea, and in Georgia, mainly in the village of Sarp, which is split by the border between Turkey and Georgia. There are also Laz villages to the East of the Sea of Marmara in Western Turkey; they emigrated there from the end of the 19th century onwards.

The number of speakers is not known. According to Andrews (1989: 176), in the 1965 Turkish census, which is the last official statistic, 26,007 people declared Laz as their mother tongue and 59,101 as their second language. On the other hand, Feurstein (1983) estimates the number of speakers at 250,000. Between 2000 and 5000 Laz live in Georgia (Holisky 1991: 397; Kutscher 2008).

The Laz population of Turkey was formerly Christian but converted to Islam in the 17th century (Toumarkine 1995: 48). The Laz are primarily agriculturists. Tea cultivation was introduced in the region in the 20th century and holds a prominent place in the local economy.

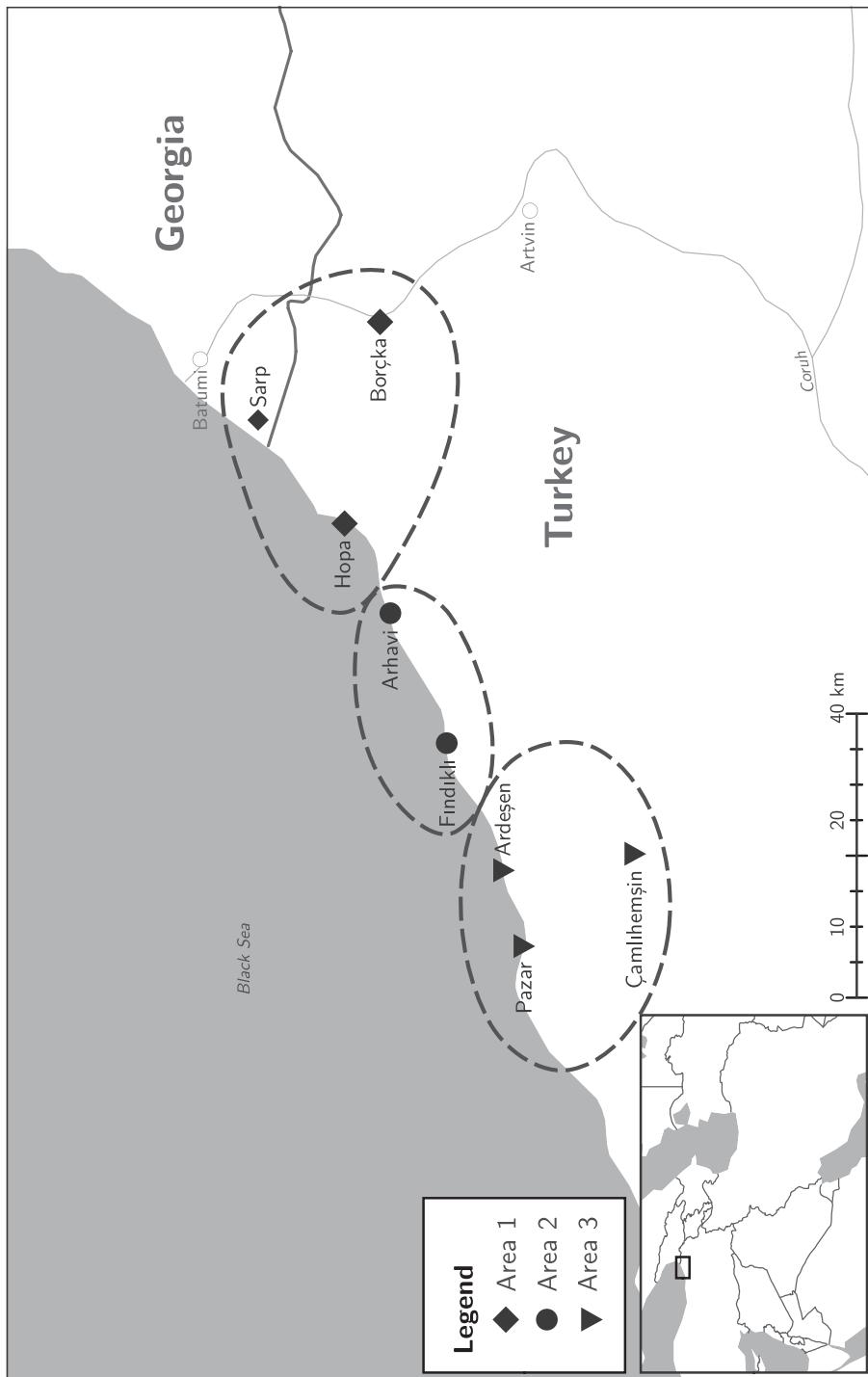


Figure 2: Map of main Laz-speaking locations

In Turkey, virtually all speakers are bilingual with Turkish. Nowadays, almost all children acquire at best a passive knowledge of the language.¹ Only among older adults does one frequently find fluent speakers. Code-switching with Turkish is widespread, especially among men. Even the speech of Laz who do not code-switch is interspersed with Turkish elements. Laz is thus a highly endangered language.

Laz was never used in school. In 2012, however, the Turkish government made the decision to offer elective Laz lessons at school, provided enough pupils ask for it.

Laz is not a written language. However, magazines written partly in Laz have been published for many years. Furthermore, the recent efforts of some Laz has led to the publication of novels in this language.

Laz has been partly described and is documented by several collections of texts (see references in Lacroix 2009a: section 1.2). In addition, a documentation project funded by HRELP (London) has been undertaken by the author.² Laz is divided into three main dialectal areas (see map). Each area is in turn divided into a number of subdialects. Mutual intelligibility between certain dialects is not straightforward, but it is not problematic between subdialects of the same area. This paper describes the Arhavi dialect.

The following sections give basic information on the phonology (§2), the noun phrase (§3), the verb (§4), simple clauses (§5) and complex clauses (§6). As it proved impossible to give even the most elementary information on many aspects of Laz in so few pages, most of the topics touched upon here were chosen either because of their typological or areal bearing.

2. Phonology

2.1. Vowels and consonants

Laz has five vowels: /a/, /e/, /i/, /o/ and /u/. Vowels /i/, /e/ and /a/ have no clearly perceptible allophonic variants. /o/ and /u/ have fronted allophones when preceded by the palatal consonant /j/ in the verbal prefixes *gj-* and *j-*. There are no diphthongs.

The consonantal system is represented in Table 1.

¹ During his one-year stay in the Laz community in 2011, the author met children who were able to speak Laz in the village of Dikkaya (North-East of Turkey) and teenagers in the village of Kabalak (West of Turkey). İrfan Çağatay (September 2012, p.c.) reports that he heard children from the village of Topluca (North-East of Turkey) speak Laz. These are exceptional cases.

² <https://elar.soas.ac.uk/Collection/MPI546814> (accessed 20 November 2017).

Table 1: Arhavi Laz consonants

		labial	dental or alveolar	post-alveol. or palatal	velar or uvular	glottal
occlusive	voiceless non-glott.	p	t		k	
	voiceless glottalized	p'	t'		k'	
	voiced	b	d		g	
affricate	voiceless non-glott.		ts	tʃ		
	voiceless glottalized		ts'	tʃ'		
	voiced		dz	dʒ		
fricative	voiceless		s	f	x	h
	voiced	v	z	ʒ	ɣ	
nasal		m	n			
lateral			l			
approximant			r	j		

Although the consonants with an apostrophe have been classified as glottalized, further research needs to be done to assess the exact nature of these consonants. Glottalization is not as strong in Laz as it is in standard Georgian. When two or more glottalized consonants follow each other, glottalization is indicated only on the last one. Thus, *metk'otfi* ‘throw it!’ holds for *met'k'otfi*. Voiceless non-glottalized stops /p/, /t/ and /k/ are aspirated. In addition to the consonants in Table 1, the phoneme /f/ is found in Turkish borrowings (e. g. *tufeyi* ‘gun’, from Turkish *tüfek*).

Laz allows up to four consonants in syllable onsets, as can be seen in the following examples:

Two consonants: *skiri* ‘child’, *ktfini* ‘old woman’

Three consonants: *tfxvapa* ‘hot’

Four consonants: *mskva* ‘beautiful’

2.2. Stress

Words other than finite verb forms are normally stressed on the penultimate syllable: *bózo* ‘girl’, *k'ap'úla* ‘back’, *okosále* ‘broom’, *ok'átfxe* ‘then’, *héya* ‘this one’. The position of the stress shifts when the word is inflected: *bozó-fi* ‘girl-GEN’, *bozo-pé-jí* ‘girl-PL-GEN’. Certain final vowels may drop as a result of free variation, which leads to stress falling on the last syllable: *dal-épe ~ dal-ép* ‘sisters’.

Stress assignment rules in finite verb forms depend on morphological and lexical properties. The details will not be given here. Suffice it to say that stress may fall on any syllable except the final one. Some examples are given in (1–3).

- (1) *b-i-gzál-are*
11-MID-leave-FUT.11/2SG
'I'll leave'
- (2) *b-o-p'aamit-áp-am*
11-CAUS-talk-CAUS-TH
'I have him talk'
- (3) *kó-dol-i-kun-es-doe*
PV-PV-MID-put_on-AOR.13.PL-EVD
'they put it on'

Negation is marked by *va(r)*, which precedes the verb immediately and forms one accentual unit with it:

- (4) *vá ge-b-ul-u*
NEG PV-11-go_down-TH
'I don't go down'

3. The noun phrase

3.1. Nominal suffixes and clitics

This section presents the formation of the plural, the possessive determiners, the cases and the “additive” suffix *-ti*. These appear to the right of the head of the NP, in the following order: plural-possessive-case-additive (5).

- (5) *džumal-epe-tʃkimi-ʃi-ti*
brother-PL-POSS1SG-GEN-ADD
'my brothers' too'

The plural is formed with the suffixes *-pe* and *-epe*, depending on phonological and lexical properties of the base: *nek'na* ‘door’, *nek'na-pe* ‘doors’; *pudʒi* ‘cow’, *pudʒ-epe* ‘cows’. The base of certain nouns ending in /a/ is expanded with *-l* in the plural: *kva* ‘stone’, *kval-epe* ‘stones’; *džuma* ‘brother’, *džumal-epe* ‘brothers’; *oxordža* ‘woman’, *oxordžal-epe* ‘women’.

The plural suffixes have both affixal and clitical characteristics. As far as affixal characteristics are concerned, the plural suffixes belong to the domain of the stress-assignment rule, as already illustrated above: *bózo* ‘girl’, *bozó-pe* ‘girl-PL’. Furthermore, the choice between *-pe* and *-epe* depends in part on lexical information, and thus cannot be accounted for entirely in phonological terms. Finally, the plural triggers allomorphic variation of the base, as it selects the base expanded with *-l*, if any.

As far as clitical characteristics are concerned, the plural suffixes can attach

not only to nouns, but also to adjectives, postpositions (6) and finite verb forms (7). The latter situation occurs in free (headless) relative clauses (see §6.1).

- (6) *tʃkimi ster-epe i-ster-t'es*
 1SG.GEN like-PL MID-play-IMPFT.13.PL
 ‘The ones like me [i. e. the children my age] used to play.’ (own data)
- (7) *si-na tʃk'om-i-pe-k*
 2SG-SUB eat-AOR-PL-ERG
 ‘those you ate’ (K72.137)

Table 2 gives the possessive determiners. Here as elsewhere in Laz, there is no grammaticalized gender distinction. Three examples for possessive determiners are provided in (8); note that they do not carry stress.

Table 2: Possessive determiners

1sg	- <i>tʃkimi</i>
2sg	- <i>skani</i>
3sg	- <i>mufi</i>
1pl	- <i>tʃkuni</i>
2pl	- <i>tkvani</i>
3pl	- <i>mutepesi</i>

- (8) *bée-tʃkuni oxórdza-tʃkimi-k ts'ip'il-épe-mufi-s*
 child-POSS1PL woman-POSS1SG-ERG young-PL-POSS3SG-DAT

Table 3 shows the seven cases of Arhavi Laz.

Table 3: Cases

absolutive	-
ergative	- <i>k</i>
dative	- <i>s</i>
genitive	- <i>fì</i>
allative	- <i>fe</i> ~ - <i>jà</i>
ablative	- <i>fen</i>
instrumental	- <i>ten</i>

Cases have both affixal and clitical characteristics. As regards affixal characteristics, cases belong to the domain of stress assignment, as already illustrated above: *bózo* ‘girl’, *bozó-fì* ‘girl-GEN’. On the morphological level, the genitive, allative and ablative cases trigger a special form of the base of 1st and 2nd person pronouns (see below).

As regards clitical characteristics, cases are generally not repeated in conjunctions:

- (9) *bozo do bit'i-s*
 girl and boy-DAT
 ‘for the girl and the boy’ (K72.139)

Furthermore, cases can attach to different types of words. In headless relative clauses, for instance, they attach to finite verb forms (10).

- (10) *ordo-na e-j-sel-asen-s gza ordo*
 early-SUB PV-MID-stand_up-FUT.13SG-DAT path early
 ‘The path of that who stands up early ends early.’ (own data)

The last element to be considered here is *-ti*, glossed ADD for “additive”, which has a range of uses. In particular, *-ti* is used with the meaning ‘and, too’ (11).

- (11) *hemu-s udži-ti k-u-y-u-n*
 DIST.DEM-DAT ear-ADD PV-II3.APPL-have-TH-13SG
toli-ti k-u-y-u-n xe-ti k-u-y-u-n
 eye-ADD PV-II3.APPL-have-TH-13SG hand-ADD PV-II3.APPL-have-TH-13SG
 ‘She has ears, and eyes, and hands too.’ (K93.130)

Second, the suffix *-ti* is used to switch from one topic to another, especially in dialogues (12).

- (12) *bere-k ma-ja hafø*
 child-ERG 1SG-RS thus
hafø m-a-yod-u-ja u-ts'-u
 thus II1-MID.APPL-happen-AOR.13SG-RS II3.APPL-tell-AOR.13SG
 ‘The young man told him: ‘Such and such things happened to me’.
- he k'otfi-k-ti si-na g-a-yod-u steri*
 DIST.DEM man-ERG-ADD 2SG-SUB II2-MID.APPL-happen-AOR.13SG like
- ma-ti hefø m-a-yod-u*
 1SG-ADD thus II1-MID.APPL-happen-AOR.13SG
 The man answered: ‘What happened to you happened to me too.
- si do ma*
 2SG and 1SG
- džumal-epe do-b-iv-a-t-ja u-ts'-u he*
 brother-PL PV-11-become-OPT-1/2PL-RS II3.APPL-tell-13SG DIST.DEM
k'otfi-k
 man-ERG
 You and me, let’s become brothers!’

<i>ha</i>	<i>bere-k-ti</i>	<i>k'aj-a</i>
PROX.DEM	child-ERG-ADD	good-RS
The young man said: 'Good'. (Ž.95)		

Here again, the element at stake has both affixal and clitical characteristics. As far as affixal characteristics are concerned, *-ti* belongs to the domain on which the stress-assigning rule obtains: *túti-k* 'bear-ERG', *tutí-k-ti* 'bear-ERG-ADD', *tut-epé-k-ti* 'bear-PL-ERG-ADD'. On the morphological level, *-ti* triggers a special singular absolute form of the demonstrative pronouns, as shown in Table 4.

Table 4: Partial paradigm of the demonstrative pronouns with and without *-ti*

		proximal demonstrative pronoun		distal demonstrative pronoun	
		without <i>-ti</i>	with <i>-ti</i>	without <i>-ti</i>	with <i>-ti</i>
singular					
absolutive	<i>haja</i>	<i>ha-ti</i>		<i>heja</i>	<i>he-ti</i>
ergative	<i>hamu-k</i>	<i>hamu-k-ti</i>		<i>hemu-k</i>	<i>hemu-k-ti</i>
dative	<i>hamu-s</i>	<i>hamu-s-ti</i>		<i>hemu-s</i>	<i>hemu-s-ti</i>

As far as clitical characteristics are concerned, *-ti* can be attached to words of different categories: nouns, pronouns, postpositions, negation *var* (see (43) below) and conditional clause subordinator *na*.

Several other languages of the area (Turkish, Zazaki and Kurmanjî Kurdish) have an enclitic particle akin to Laz *-ti* used as a topic-switch marker (Haig 2001: 207), and similar items are also found in the languages of northern Iraq, for example North Eastern Neo-Aramaic (Khan, this volume, chapter 3.4) and Kurdish (Haig, this volume, chapter 3.3).

3.2. Pronouns, demonstratives and interrogatives

Table 5 gives the paradigm of 1st and 2nd person pronouns, which involves some variation. 3rd person pronouns are identical to demonstratives, and are presented further below. As can be seen, 1st and 2nd person pronouns have the same form in the three cases indicating core syntactic roles: absolute, ergative and dative. Their alignment is thus neutral (A=O=S).

Table 5: Declension of 1st and 2nd person pronouns

	1sg	2sg	1pl	2pl
absolutive				
ergative	<i>ma ~ man</i>	<i>si ~ sin</i>	<i>tʃku ~ tʃkun</i>	<i>tkva ~ tkvan</i>
dative				
genitive	<i>tʃimi</i>	<i>skani</i>	<i>tʃuni</i>	<i>tkvani</i>
allative	<i>tʃim-de ~</i> <i>tʃim-da</i>	<i>skan-de ~</i> <i>skan-da</i>	<i>tʃun-de ~</i> <i>tʃun-da</i>	<i>tkvan-de ~</i> <i>tkvan-da</i>
ablative	<i>tʃim-den ~</i> <i>tʃimi-sen</i>	<i>skan-den ~</i> <i>skani-sen</i>	<i>tʃun-den ~</i> <i>tʃuni-sen</i>	<i>tkvan-den ~</i> <i>tkvani-sen</i>

Laz has proximal and distal deictics. Only the demonstrative pronouns and determiners are considered here. The demonstrative pronouns have two bases: one in the absolutive (proximal *haja*, distal *heja*), the other in all remaining cases (proximal *hamu-*, distal *hemu-*). The demonstrative pronouns take the same case suffixes as nouns (see the partial paradigms in Table 4). Proximal and distal demonstratives are used as 3rd person pronouns, as can be seen in examples (13) and (14), respectively.

- (13) *ma-ja hamu-s a muntxa b-u-ts'v-a-ja*
 1SG-RS PROX.DEM-DAT one something 11-II3.APPL-tell-OPT-RS
 [A man, a jackal and a snake are talking to each other. Referring to the man, the jackal says to the snake:] ‘I’ll say something to him.’ (own data)

- (14) *hemtepe k'ala avi-se b-id-i favsati-se*
 DIST.DEM.PL with hunting-ALL 11-go-AOR Şavşat-ALL
 ‘[Last year, six men came, six Europeans.] I went hunting with them in Şavşat.’ (own data)

The demonstrative determiners are *ha ~ ham* (proximal) and *he ~ hem* (distal). They do not inflect. Examples can be found in (12) and (54).

Table 6 gives a list of interrogative proforms.

Table 6: Interrogative proforms

<i>mi</i>	‘who?’
<i>mu</i>	‘what?’
<i>so</i>	‘where?’ (location and direction)
<i>solen ~ sole</i>	‘from where? through where?’
<i>mudes ~ mundes</i>	‘when?’
<i>mutʃ'o</i>	‘how?’
<i>nak'o</i>	‘how much? how many?’

Mi ‘who’ and *mu* ‘what’ decline as nouns and can be pluralized (*mipe*, *mupe*). Interrogative proforms generally occur in focus position, i. e. in front of the verb (see §5.1), as in example (15).

- (15) *ma mundes b-yur-are*
 1SG when 11-die-FUT.11/2SG
 ‘When will I die?’ (D67.XXVIII)

3.3. Adjective

Laz adjectives do not inflect. They occur on the left of the noun they determine (see §5.1). Some underived adjectives are given in Table 7.

Table 7: Sample of underived adjectives

<i>didi</i>	‘big, great’	<i>subuki</i>	‘light (of weight)’
<i>tʃ'ut'a</i>	‘small, little, tiny’	<i>mtʃ'ipe</i>	‘slender, thin’
<i>t'abala</i>	‘flat’	<i>mtʃxu</i>	‘thick’
<i>mskva</i>	‘beautiful, handsome’	<i>gundze</i>	‘long’
<i>k'ai</i>	‘good’	<i>mk'ule</i>	‘short’
<i>p'at'i</i>	‘bad, ugly’	<i>ayne</i>	‘new, young’
<i>utʃa</i>	‘black’	<i>mdʒveʃi</i>	‘old’
<i>ktʃe</i>	‘white’	<i>t'utsa</i>	‘hot’
<i>mtʃ'ita</i>	‘red’	<i>mendra</i>	‘far away’
<i>monk'a</i>	‘heavy’		

3.4. Echoic

Laz has a productive derivational mechanism which consists of reduplicating a word and modifying the reduplicant according to the following rule: when the word begins with a vowel, /m/ is added in front of it (16); when it begins with one or several consonants, these are replaced by /m/ (17). The resulting compound means ‘x and other things related to x’ (18).

- (16) *ejer* ‘saddle’ (Turkish loan) → *ejer-mejer*
optʃk'omi ‘I ate’ → *optʃk'omi-moptʃk'omi*
 (elicited example)
- (17) *dadzi* ‘thorns’ → *dadzi-madzi*
tsxeni ‘horse’ → *tsxeni-menı*

- (18) *xortsi-mortsi var*
 meat-ECHO NEG
xvala mkvei tf'v-er mef-o-by-u-n dolox
 only flour burn-PART PV-APPL-be-TH-13SG inside
 '[People cook these ingredients and call it *kete*.] There's no meat or
 anything like that in it, but only cooked flour.' (K93.99)

This morphological process, called “echoic” by Haig (2001), is an areal phenomenon, attested in Turkish, Kurmandjî Kurdish, Zazaki, Persian, Armenian and Georgian, as well as in languages of the Balkans (Haig 2001: 208–209). A similar phenomenon is attested in the Indian linguistic area (Emeneau 1980: 114).

4. The verb

The morphology of Laz finite verbs encodes tense, aspect, mood, evidentiality, person, valency and locative distinctions. Four prefixes, called “affirmative preverbs”, have a range of additional functions (see below). Table 8 shows the order of the verbal morphemes.

Table 8: Morphological slots in Laz finite verb forms

-4	preverb
-3	preverb
-2	person marking
-1	valency/TAM
0	root
1	causative
2	causative
3	thematic suffix/TAM
4	TAM
5	TAM
6	person marking
7	TAM/evidentiality

Morphologically, Laz verbs may be divided into two broad classes, on the basis of the suffix realizing the feature “Set I, 3sg” in the present tense: Class 1 verbs have *-s*, while Class 2 verbs have *-n*. Thematic suffixes as well differ. These suffixes, glossed “TH”, occur in certain tenses only, like the present (see 19–20) and the imperfect (see 27); in other tenses, like the aorist and the imperative, no thematic suffix appears (see 22 and 63, respectively). Class 1 verbs have either no thematic suffix or one of the thematic suffixes *-am*, *-em*, *-im*, *-om* and *-mer*; Class 2

verbs have one of the thematic suffixes *-er*, *-ir* or *-ur*³. These morphological distinctions correlate with syntactic and semantic distinctions. In particular, all Class 2 verbs are intransitive.

The following sections present person marking, preverbs, tenses, valency-changing derivations and non-finite verb forms.

4.1. Person marking

Laz verbs are specified for at least one person-number feature and at most two person-number features. These features will be called “Set I” and “Set II”. They are both realized through prefixes and suffixes (slots -2 and 6 in Table 8). In transitive verbs, the subject triggers agreement by means of Set I features, while the object triggers agreement by means of Set II features. Thus in (19), the subject *bozok* ‘girl’ is represented in the verb by the suffix *-s* “Set I 3sg”, while the object is represented by the prefix *m-* “Set II 1st person”. In the intransitive verb illustrated in (20), the subject *dʒuma-tʃkimi* ‘my brother’ triggers agreement by means of Set I features, realized in this case by the suffix *-n* “Set I 3sg”. An argument which triggers agreement by Set I features will be called “Set I argument”, and an argument which triggers Set II features in the verb will be called “Set II argument”.

- (19) *bozo-k* *ma* *m-dzi-om-s*
 girl-ERG 1SG 111-see-TH-13SG
 ‘The girl sees me.’ (own data)
- (20) *dʒuma-tʃkimi* *tk'ob-u-n*
 brother-POSS1SG hide-TH-13SG
 ‘My brother hides himself.’ (own data)

Table 9 gives the paradigm of affixes realizing Set I features, and Table 10 the paradigm of affixes realizing simultaneously Set I and Set II features. No verb form is specified for Set II features only.⁴ In both Tables 9 and 10, the choice between *-s*, *-n* and *-u* on the one hand and between *-an*, *-nan*, *-es* and *-n* on the other is conditioned by the morphological class to which the verb belongs and by tense. The dash indicates the position of the stem.

³ These thematic suffixes are realized as *-e*, *-i* and *-u*, respectively, when followed by a person suffix beginning with /n/.

⁴ Lacroix (2014) puts forward a historical scenario explaining the origin of the distribution of person-number suffixes in Table 10, the organization of which is difficult to account for in synchrony.

Table 9: Set I paradigm

	Singular	Plural
1	b—	b—t
2	—	—t
3	—s/n/u	—an/nan/es/n

Table 10: Sets I-II paradigm

	1sg	1pl	2sg	2pl	3sg	3pl
1sg			g—		b—	
1pl			g—t			b—t
2sg	m—	m—t			—	
2pl					—t	
3sg	m—s/n/u/	m—an/nan/es/n	g—s/n/u		—s/n/u	—an/nan/es/n
3pl			g—an/nan/es/n		—an/nan/es/n	

In a clause containing a 3rd person Set I argument and a 3rd person Set II argument, only one of them can trigger plural agreement. In the transitive construction, the Set I argument has this ability (see *bozopek* ‘girls’ and the plural suffix *-an* in 21a), while the Set II argument does not (see *bitf’epē* ‘boys’ in 21b).

- (21) a. *bozo-pe-k* *bitf’i* *dzi-om-an*
 girl-PL-ERG boy see-TH-13.PL
 ‘The girls see the boy.’ (own data)
- b. *bozo-k* *bitf’-epē* *dzi-om-s*
 girl-ERG boy-PL see-TH-13SG
 ‘The girl sees the boys.’ (own data)

Consider now examples (22a–c), which illustrate the potential derivation.⁵ The argument referring to the participant who can do the action triggers agreement by Set II features (cf. *m-* “Set II 1st person” in 22a); it is marked by the dative case (cf. *k’otsepēs* ‘men’ in 22b and *k’otfis* ‘man’ in 22c). This argument has the ability to trigger plural agreement (see the suffix *-es* in 22b), while the Set I argument does not (see *tʃomepe* ‘fish’ in 22c).

⁵ This derivation is called after its most frequent reading; it has other uses as well (see §4.4).

- (22) a. *tfxomi va m-a-tʃ'op-u*
 fish NEG III-POT-take-AOR.13SG
 ‘I could not catch fish.’ (own data)
- b. *k'otʃ-epe-s tfxomi va a-tʃ'op-es*
 man-PL-DAT fish NEG POT-take-AOR.13.PL
 ‘The men could not catch fish.’ (own data)
- c. *k'otʃi-s tfxom-epe va a-tʃ'op-u*
 man-DAT fish-PL NEG POT-take-AOR.13SG
 ‘The man could not catch the fish.’ (own data)

The term “inverse construction” will be used to characterize any construction like the potential, where a Set II 3rd person argument has the ability to trigger plural agreement. The term “plain construction”, in contrast, will characterize any construction in which a Set I 3rd person argument has the ability to trigger plural agreement. The inverse construction is found in four situations: the potential derivation; the benefactive derivation, which adds a possessor, beneficiary or maleficiary to the valency of the verb; the perfect tenses (perfect, pluperfect II and evidential pluperfect II); and lexicalized verbs and idioms involving an experiencer or a possessor. All finite verbs in Laz are specified for Set I. One-place verbs in the inverse construction include a Set I 3rd person marker which does not refer to any argument (see among others 22a).

In all these constructions, the Set II argument is in the dative case. As it exhibits several subject properties (see Lacroix 2009a, section 11.2), it can be considered as a non-canonical subject.

The term “inverse construction” (or “inversion”) is common in South Caucasian linguistics. It conveys the idea that the verbal affixes, which represent the subject and the complement of a plain construction, are “switched” in the inverse construction, where they represent the complement and the subject, respectively. A similar phenomenon can be observed in two other languages of the area, Kurdish and Aramaic (see Haig and Khan, this volume, chapter 6.1 for discussion).

4.2. Preverbs

There are two types of preverbs in Laz: affirmative and locative. Affirmative preverbs appear in slots -4 and -3. They are related to several characteristics of the clause. In particular, they never occur in clauses negated with the standard negation *var* (hence their name), nor in relative clauses. When several clauses are coordinated, affirmative preverbs tend to occur only on the verb of the last conjunct (cf. *d-* in 23).

- (23) *id-es id-es do d-u-lumdʒ-es*
 go-AOR.I3.PL go-AOR.I3.PL and PV-II3.APPL-get_dark-AOR.I3.PL
 ‘They walked, walked, and got overtaken by night.’ (K93.120)

Affirmative preverbs seem also to be sensitive to information structure. In (24), for instance, where the object is focused, the verb cannot take the affirmative preverb *-o*.

- (24) *ukui p-tʃk'om-i *optʃk'omi*
 apple 11-eat-AOR
 [What did you eat, an apple or a pear?] ‘I ate an apple.’ (own data)

In addition, affirmative preverbs are used for the flexion of certain tenses (see §4.3). Affirmative preverbs are one of the most intricate issues in the grammar of Laz. Further research is needed to figure out all the rules which determine their use.

There are more than thirty locative preverbs, all of which occur in slot -3. Morphologically, these can be simple (*me-*, *mo-*, *e-*, *ge-*, etc.) or complex (*mesa-*, *mota-*, *mek'a-*, *mok'a-*, etc.). A few examples are given in (25).

- (25) *mo-xt-u* ‘he came’
 ama-xt-u ‘he came in’
 gama-xt-u ‘he went out’
 e-xt-u ‘he went up’
 ge-xt-u ‘he went down’
 mesa-xt-u ‘he went (into a narrow place)’
 dolo-xt-u ‘he went down (into a narrow place)’

Locative preverbs are not restricted to motion verbs, as can be seen in (26).

- (26) *ma sanduyi-s dolo-p-x-er*
 1SG coffer-DAT PV-11-sit-TH
 ‘I’m sitting in the coffer.’ (Q.48)

Locative preverbs may be lexicalized. For instance, the following verbs can be derived on the root *-tʃ-*: *-tʃ-* (without preverb) ‘feed’, *gama-tʃ-* ‘sell’, *me-tʃ-* and *mo-tʃ-* ‘give’, *ge-tʃ-* ‘hit’, *gela-tʃ-* ‘play [an instrument]’.

4.3. Tenses and evidentiality

Tenses have exponents in various slots in the verb, as can be seen in Table 8. A list of Arhavi Laz tenses is given in Table 11. Tenses in column 2 are formed on those of column 1 by adding an affirmative preverb; tenses in column 3 are formed on those of column 1 by adding the evidential suffix *-doren* (or one of its phonologically reduced variants); the general evidential imperfect (column 4) is formed by

adding both an affirmative preverb and the evidential suffix to the imperfect. This is illustrated in (27). The construction of verbs in the perfect, pluperfect II and evidential pluperfect II tenses is inverse (see §4.1).

Table 11: Tenses

1	2	3	4
present	general present	evidential present	
imperfect	general imperfect	evidential imperfect	general evidential imperfect
aorist		evidential aorist	
pluperfect I			
subjunctive			
optative			
past optative			
imperative			
future	hypothetical future		
perfect			
pluperfect II		evidential pluperfect II	

- (27) imperfect general imperfect
tʃ'ar-um-t'i *do-tʃ'ar-um-t'i*
 write-TH-IMPFT PV-write-TH-IMPFT
 ‘you were writing’ ‘you used to write’
- evidential imperfect general evidential imperfect
tʃ'ar-um-t'i-doe *do-tʃ'ar-um-t'i-doe*
 write-TH-IMPFT-EVD PV-write-TH-IMPFT-EVD
 ‘you were writing, they say’ ‘you used to write, they say’

The evidential suffix *-doren* (or one of its variants) occurs in the right-most slot in finite verb forms. It is used when reporting events related by another person, which the speaker did not witness:

- (28) *bere-t/kimi-k yali-k u/kuri*
 child-POSS1SG-ERG river-ERG apple
gel-i-Ø-me-t'u-fi ar teyi ko-dzir-u-doren
 PV-MID-carry_along-TH-IMPFT.13SG-TEMP one unit PV-see-AOR.13SG-EVD
 [A child finds an apple in a river and shows it to his father. The father asks: 'Where did you find this apple?' The son answers: 'I caught it as the river carried it along.' The father decides to find the owner of the apple. He says to a man he meets:] 'My son saw an apple which was carried along by the river.' (Ž.27)

Further, the evidential is used when drawing an inference based on present evidence:

- (29) *xodza do-yur-u-doren*
 hodja PV-die-AOR.13SG-EVD
 [A group of men see a hodja lying motionless on the ground. They say:] 'The hodja has died!' (D67.XXVIII)

Finally, verb forms with *-doren* may have a mirative meaning:

- (30) *haja nak'o i-monk'an-u-doren*
 PROX.DEM how_much MID-get_heavy-AOR.13SG-EVD
 [An old man carries another man he has captured in a basket. At some point, he puts the basket on the ground and goes away for a while. In his absence, the captured man slips out of the basket, fills it with stones and manure, binds it again and goes away. The old man comes back. As he wants to put the basket on his back again, he says:] 'How much heavy this has gotten!' (Ž.42)

4.4. Valency-changing derivations

Laz, as the other South Caucasian languages, has several valency-changing derivations. These have incidences on various parts of verbal morphology: valency markers (slot -1), causative suffixes (slots 1 and 2), thematic suffixes (slot 3), 3rd person singular and plural suffixes (slot 6) and choice between the plain and the inverse construction are affected by these derivations. The main derivations will be illustrated here with verbs formed on the root *-by-* and the locative preverb *dolo-*. Verbs formed on the root *-by-* are used when talking about a group of similar countable objects (apples, blades of grass, lice, flies, sheep, people, djinns, etc.) or of an amount of matter (ash, plaster, etc). The preverb *dolo-* indicates a downward movement into a narrow place or position in such a place. For each verb, the case of the arguments, the Set of person-number features they trigger in the verb and possible adjuncts are indicated in angle brackets. The argument which has the

ability to trigger plural agreement is typed in bold. The Class of the verb is given in parentheses.

The first example, (31), contains a plain transitive verb; its valency marker (slot -1) is *o-*. The presence of *o-* in plain transitive verbs is not predictable and is specified in the lexicon (the verb ‘see’, for instance, doesn’t take this prefix (see 19)).

- (31) plain transitive ‘<**erg I**> tips <**abs II**>’ (Class 1)

<i>Xalili-k</i>	<i>uskui</i>	<i>dol-o-by-am-s</i>
Halil-ERG	apple	PV-TR-tip-TH-13SG
‘Halil tips the apples.’ (own data)		

The following two examples illustrate the middle derivation. Middle verbs are marked by *i-* in slot -1. They may belong to Class 1 or Class 2. The verb in example (32) is transitive and belongs to Class 1. It indicates that the subject is directly affected or concerned by the event described by the verb (in this case, the subject is the possessor of the adjunct *dzebis* ‘pocket’ and is thus both the initiator and the goal of the event). Class 1 middle verbs may have other readings.

- (32) transitive middle ‘<**erg I**> tips <**abs II**> into his <**adjunct**>’ (Class 1)

<i>Xalili-k</i>	<i>uskui</i>	<i>dzebi-s</i>	<i>dol-i-by-am-s</i>
Halil-ERG	apple	pocket-DAT	PV-MID-tip-TH-13SG
‘Halil tips the apples in his pocket.’ (own data)			

The middle verb in (33) belongs to Class 2. The reading of this intransitive verb is anticausative, but here again, other readings exist. As can be seen, the thematic suffix (-*e*) and the Set I 3sg suffix (-*n*) differ from those of the Class 1 verb in (32). Middle voice in Laz is presented in more details in Lacroix (2012a).

- (33) intransitive middle ‘<**abs I**> tips’ (Class 2)

<i>uskui</i>	<i>dol-i-by-e-n</i>
apple	PV-MID-tip-TH-13SG
‘The apples scatter.’ (own data)	

While the middle derivation illustrated in (33) is valency-decreasing, the applicative, of which (34) is an example, adds one argument to the valency of the corresponding plain transitive, making it ditransitive. The applicative argument (*Xalili-s* in this example) is in the dative and triggers Set II agreement.⁶ It has the semantic role of beneficiary. Other possible semantic roles are maleficiary, possessor, location, goal and addressee. The valency prefix of this derivation is *i-* when the applicative argument is 1st or 2nd person and *u-* when it is 3rd person.

⁶ Case marking of the applicative argument is not the same as that of the object, which is in the absolute. In consequence, this derivation should better be analyzed as a ‘non-canonical’ applicative.

- (34) transitive applicative ‘<erg I> tips <abs> for <dat II>’ (Class 1)

Xalili-s u/kui dolo-b-u-by-am
 Halil-DAT apple PV-11-II3.APPL-tip-TH
 ‘I’m tipping apples for Halil.’ (own data)

The benefactive derivation, illustrated in (35), is marked by *a* in the preroot slot. The construction includes two arguments, one in the dative and one in the absolute. The dative argument may have the semantic role of beneficiary, possessor or maleficiary. It shares many properties with an applicative argument, but differs from it in one important respect: it has the ability to trigger plural agreement. The dative argument here is a non-canonical subject.

- (35) benefactive ‘<dat II>’s <abs I> tips’ (Class 2)

Xalili-s u/kui dol-a-by-e-n
 Halil-DAT apple PV-MID.APPL-tip-TH-13SG
 ‘Halil’s apples scatter.’ (own data)

The same prefix *a* appears in the potential derivation (36), already discussed above. As mentioned there, this derivation has other readings. In particular, it is used to indicate that the subject acts accidentally and involuntarily (37).⁷

- (36) potential ‘<dat II> is able to tip <abs I>’ (Class 2)

Xalili-s u/kui dol-a-by-e-n
 Halil-DAT apple PV-POT-tip-TH-13SG
 ‘Halil is able to tip the apples.’ (own data)

- (37) deagentive ‘<dat II> tips <abs I> accidentally’ (Class 2)

Xalili-s u/kui ko-dol-a-by-u
 Halil-DAT apple PV-PV-POT-tip-AOR.13SG
 ‘Halil tipped the apples inadvertently.’ (own data)

The benefactive derivation (35) and the potential/deagentive derivation (36–37) can be distinguished from one another on the basis of semantic and syntactic properties. In the benefactive derivation, the referent of the dative argument is affected by the event described by the verb, but is not its instigator. It cannot be the addressee of a command. By contrast, the referent of the dative argument in the potential/deagentive derivation is the instigator or source of the process. In the case of the deagentive, the dative argument can be the addressee of an order (38).

- (38) *tabay-epē mo me-g-a-tk’ot/-ap-u-t’as*

plate-PL PROH PV-112-POT-throw-CAUS-TH-SUBJ-13SG
 ‘[Throw away everything, but watch out!] Don’t throw away the plates!’
 (own data)

⁷ This “deagentive” reading seems more natural in the aorist tense than in the present.

Historically, both the benefactive and the potential/deagentive derivations result from a middle-applicative derivation; the dative argument was an applicative argument, which took on subject properties in the course of time. This explains the morphological and syntactic similarities between these constructions. The history of the potential/deagentive derivation is discussed in Lacroix (2009a, section 9.8.6.2.3).

The causative derivation is marked by the suffixes *-ap* or *-in* immediately after the root. *-ap* is used when deriving a causative from a transitive or intransitive verb (39), while *-in* is used only for causatives derived from intransitive verbs. Causative verbs take the preroot prefix *o-*.

- (39) causative ‘<**erg I**> has <**dat II**> tip <**abs**>’ (Class 1)

<i>Xalili-k</i>	<i>Ali-s</i>	<i>ufkui</i>	<i>dol-o-by-ap-am-s</i>
Halil-ERG	Ali-DAT	apple	PV-CAUS-tip-CAUS-TH-13SG
‘Xalil has Ali scatter apples.’ (own data)			

The same root *-by-* may be used to derive the applicative stative verbs ‘x is somewhere’ and ‘somebody’s x is somewhere, somebody has x somewhere’, where x is a group of similar countable objects or an amount of matter. These verbs are characterized by the thematic suffix *-u(r)*. They are intransitive. The first one is illustrated in (40). Its valency marker is *o-*⁸. The applicative argument *t’ik’inas* ‘basket’ indicates the place where the apples are located. The second verb (41) co-occurs with two dative NPs. The first one, *Xalilis*, has the semantic role of possessor. The second dative NP, *t’ik’inas*, is an adjunct and indicates the place where the apples are located. The verb takes the valency marker *i-* when the Set II argument is 1st or 2nd person and *u-* when it is 3rd person. Historically, this construction derives from an applicative construction; hence the similarities between them.

- (40) intransitive locative applicative ‘<**abs I**> is in <**dat II**>’ (Class 2)

<i>ufkui</i>	<i>t’ik’ina-s</i>	<i>dol-o-by-u-n</i>
apple	basket-DAT	PV-APPL-be-TH-13SG
‘There are apples in the basket.’ (own data)		

- (41) intransitive possessive applicative ‘<**dat II**> has <**abs I**> in <**adjunct**>’ (Class 2)

<i>Xalili-s</i>	<i>t’ik’ina-s</i>	<i>ufkui</i>	<i>dol-u-by-u-n</i>
Halil-DAT	basket-DAT	apple	PV-II3.APPL-tip-TH-13SG
‘Halil has apples in his basket.’ (own data)			

⁸ As is clear by now, the prefix *o-* in slot -1 has a range of functions: it appears in certain transitive verbs (31), in all causative verbs (39), and in locative applicative verbs (40).

4.5. Negation

As we saw in section 2.2, negation is expressed by *var*, which is stressed and precedes the verb. All pre-root slots can be filled in negative verbs except slot 4, which hosts the affirmative preverb *ko-* (see §4.2); this is illustrated in (42). The additive suffix *-ti* can be added to the negation, yielding the meaning ‘even not’ (43), which suggests that *var* is not completely bound to the verb.

- (42) *var do-m-i-dʒox-i*
 NEG PV-II1-APPL-call-AOR
 ‘you didn’t call me’ (D37.V)

- (43) *var-ti ox-a-nk’an-es-doren*
 NEG-ADD PV-POT-wobble-AOR.13.PL-EVD
 ‘They could even not manage to wobble him.’ (D67.I)

4.6. Non-finite verbs

Laz has four non-finite verb derivations: verbal noun, general participle, future participle and negative participle. The morphological slots of non-finite verb forms are presented in Table 12.

Table 12: Morphological slots in non-finite verb forms

-2	<i>u-</i> (negative participle)
-1	preverb
0	root
1	derivational suffix (<i>-am</i> , <i>-im</i> , <i>-um</i> , <i>-in</i> , <i>-ap</i> , <i>-al</i> , <i>-mal</i>)
2	<i>-u</i> (verbal noun, negative participle), <i>-eri</i> (general participle), <i>-oni</i> (future participle)

There is some variation in the choice of the derivational suffix in slot 1. The future participle of the verb ‘see’, for instance, can be *u-dzir-am-u* (52) or *u-dzir-ap-u* (53).

The verbal noun inflects like a noun and can appear in the same positions as nouns. The internal structure of verbal noun constructions involves some variation: the object of the corresponding transitive finite verb may appear in the genitive or in the absolute. Both cases are illustrated in sentence (44), where *k’andyu* ‘strawberry’ is in the absolute, while *oda* ‘room’ is in the genitive.

- (44) *ma k’andyu o-gor-u seni mo-p-t-i*
 1SG strawberry PV-look_for-VN for PV-11-come-AOR
oda-fi o-kos-u seni var mo-p-t-i
 room-GEN PV-wipe-VN for NEG PV-11-come-AOR
 ‘I came here to get strawberries, not to clean a room.’ (D67.IX)

The general participle may be used in an attributive, predicative (45–46) or adverbial function (47). The general participle is not syntactically oriented towards any particular argument. For instance, it determines an argument representing the object of the corresponding transitive verb in (45), an argument representing the subject of the corresponding transitive verb in (46) and an argument representing the subject of the corresponding intransitive verb in (47)⁹. Furthermore, the general participle may have a perfective (45–46) or imperfective meaning (47).

- (45) *uʃkuri tʃk'om-eri t'u*
apple eat-PART be.IMPFT.13SG
'The apples were eaten.' (state) (Ž.4)
- (46) *ma gjari tʃk'om-eri b-ore*
1SG food eat-PART 11-be
'I've already eaten.' (Ž.89)
- (47) *t'orodzi bozo-ʃe putx-in-ej ko-mo-xt-u*
dove girl-ALL fly-SUF-PART PV-PV-come-AOR.13SG
'The dove flew towards the girl.' (D67.V)

The future participle can be used as an attribute (48) or in predicative function (49).

- (48) *o-tʃaliʃ-oni jer-epə var m-a-dzir-u-ja*
PV-work-FP place-PL NEG II1-POT-see-AOR.13SG-RS
'I couldn't find a place to work.' (Ž.38)
- (49) *ʒin k'at'i o-kos-oni m-i-y-u-nan*
above floor PV-sweep-FP II1-APPL-have-TH-13.PL
'We have to sweep the above floor.' (D67.IX)

The negative participle can be used as an attribute (50), in predicative (51) or adverbial function (52). Here again, the internal structure of negative participle constructions involves some variation: in (52), the object of the corresponding finite verb, *dunja*, is in the absolute, while in (53), it is in the genitive case (*arti-katifi*).

- (50) *ar tʃ'ut'a u-gub-u xortsi*
one little NP-COOK-NP meat
'some uncooked meat' (D67.I)
- (51) *masuma-na t'u u-tʃil-u t'u*
third-SUB be.IMPFT.13SG NP-marry-NP be.IMPFT.13SG
'The third one was not married.' (K72.137)

⁹ The suffix *-ej* in example (47) is a free variant of *-eri*.

- (52) *dunja u-dzir-am-u hefo i-k'itx-om-t'es-doren*
 world NP-see-SUF-NP thus MID-study-TH-IMPFT.13.PL-EVD
 '[The school was under the earth.] They studied that way, without seeing
 the world.' (D67.I)
- (53) *xut ts'ana-s artikati-f u-dzir-ap-u orop-er-ep*
 five year-DAT RECIPR-GEN NP-see-SUF-NP love-PART-PL
 'the lovers, who hadn't seen each other for five years' (D37.VIII)

5. Simple clauses

5.1. Word order

Basic constituent order is SOV (see, among others, exs. 19, 21 and 31). Depending on information structure, other orders are possible. The main regularity is that topicalized terms are fronted while focalized terms occur in immediate preverbal position. In addition, elements known from preceding context can follow the verb. Examples can be seen in (66) and (72) below, where the postposed dative arguments are already known from the preceding context.

Arhavi Laz has postpositions and no prepositions (examples can be found in (54), (71a) and (73)). Demonstratives, adjectives, genitival complements, numerals and, most often, relative clauses precede the head nominal (54–55).¹⁰

- (54) *he mskva bozo seni*
 DIST.DEM beautiful girl for
 'for that beautiful girl' (own data)
- (55) *Nurteni-fi sum bee*
 Nurten-GEN three child
 'Nurten's three children' (own data)

5.2. Alignment

The subject of an intransitive verb may be in the absolute, ergative or dative case. *Ergative* intransitive subjects are animate and in many cases have control over the event described by the verb. *Dative* intransitive subjects occur in inverse constructions (see §4.1). *Absolute* subject intransitive verbs constitute the larger class. The alignment of these verbs is mixed: case marking follows an ergative pattern (O and S are in the absolute and A is in the ergative) while verb agreement follows an accusative pattern (A and S trigger Set I features, while O triggers

¹⁰ Relative clauses are discussed in the next section.

Set II features). Examples of absolutive subject intransitive verbs can be found in (20) and (29); for transitive verbs, see (19) and (31).

5.3. Adjuncts

Adjuncts may be in any case except the ergative. Only the dative, which has a wide range of functions, will be illustrated here. This case marks different spatial roles (location (41), goal (56) and source (57)), different time roles (location (58), duration (59), frequency (60) and time necessary for the fulfillment of an action (61)) as well as price (62).¹¹

- (56) *lazi-ſi oxori-s ar k'intſi mo-xt-u*
 Laz-GEN house-DAT one bird PV-come-AOR.13SG
 ‘A bird came to the Laz’ house.’ (Ž.30)
- (57) *Axmet' jemluyi-s ko-gama-xt-u*
 Ahmet manger-DAT PV-PV-go_out-AOR.13SG
 ‘Ahmet went out of the manger.’ (D37.XII)
- (58) *k'iſi-s termoni b-i-kom-t'i-t*
 winter-DAT termoni 11-MID-do-IMPFT-PL
 ‘In winter we cooked *termoni* [a dish].’ (own data)
- (59) *sum dya do sum seri-s tʃ'anda t'u*
 three day and three night-DAT reception be.IMPFT.13SG
 ‘The reception lasted three days and three nights.’ (Ž.35)
- (60) *k'at'a ts'ana-s mtvii mt-um-s*
 each year-DAT snow snow-TH-13SG
 ‘It snows every year.’ (own data)
- (61) *sum dek'ik'e-s azlijə-f saraji-fe ko-mo-xt-u-doren*
 three minute-DAT dragon-GEN palace-ALL PV-PV-come-AOR.13SG-EVD
 ‘He came to the dragon’s palace in three minutes.’ (D67.I)
- (62) *haa xut lia-s e-p-tʃ'op-i*
 PROX.DEM five lira-DAT PV-11-buy-AOR
 ‘I bought it for five liras.’ (own data)

¹¹ In addition to these uses, the dative marks indirect objects (as in the applicative derivation – §4.4), non-canonical subjects (in the inverse construction – §4.1), and the complement of certain postpositions.

6. Complex clauses

6.1. Relative clauses

As already mentioned, relative clauses precede the head nominal.¹² Their main verb is finite in the sense that the same form can head independent clauses. It occurs at the end of the relative clause. The relative clause includes at least one occurrence of the clitic subordinator *na*. Relative clauses use a gap strategy: the position relativized is left empty. All these properties are illustrated in example (63).

- (63) *hats'i si [divi-na dʒan-s jer] do-m-o-gur-i*
 now 2SG giant-SUB lie-13SG place PV-II1-TR-teach-IMP
 ‘Now tell me the place where the giant is lying.’ (D37.VIII)

Subjects, direct objects, indirect objects and obliques may be relativized, using the same strategy.

In free relative clauses (i. e. relatives lacking a head nominal), the plural and case markers, which normally occur on the head of the NP, are directly attached to the verb of the relative. Examples can be found in (7) and (10) above.

As we are going to see, the clitic subordinator *na* appears in other types of complex clauses. In the speech of some informants, it is attached to the following word, while other informants attach it to the preceding word in the dependent clause. The subordinator *na* may occur more than once in the dependent clause and may be inserted in an NP which is itself embedded in the relative clause (64). When repeating a sentence containing *na*, my consultant often changes its position in the clause, which suggests that it is not determined by rigid phonological or syntactic rules. The subordinator *na* is examined more exhaustively in Lacroix (2012b).

- (64) *he-na ktfini-na dolo-tk'otf-es oxordʒa do k'otfi*
 DIST.DEM-SUB old_woman-SUB PV-throw-AOR.13.PL woman and
 man
 ‘the woman and the man who threw that old woman’ (own data)

The relativizing strategy of Laz is typologically and areally uncommon. Broadly speaking, Turkic and North Caucasian languages, which mainly use prenominal participial relatives, contrast with Indo-European and South Caucasian languages, where postnominal relatives headed by finite verb forms are common. In Laz, relative clauses are prenominal and headed by finite verb forms. This strategy

¹² The corpus of published Arhavi Laz texts includes a few occurrences of another relativization strategy, where the relative is postnominal (see Lacroix 2009a, section 12.2.6).

thus differs from those found in the other South Caucasian languages and in other neighboring languages (Lacroix 2009b).

6.2. Complement clauses

Verbal complement clauses use several strategies. In (65), the internal structure of the complement clause is identical to that of relative clauses (see the preceding section).

- (65) *bere-k [mgej-epē-na t'u] k-ox-o-ts 'on-u*
 child-ERG wolf-PL-SUB be.IMPFT.13SG PV-PV-TR-understand-AOR.13SG
 ‘The child understood that it was the wolves.’ (D67.II)

Other verbs using this strategy are *-dzir-* ‘see’, *-gn-* ‘hear’, *-n-* ‘want’, *-ts 'v-* ‘tell’ and *-tʃk-* ‘know’.

A second type of complement clause makes use of the complementizer *ki*, which originates in Persian and is also used in Turkish, Georgian and several other languages of the area. Complement clauses with *ki* occur after the matrix verb. Phonologically, the complementizer attaches to the last word of the matrix clause (66).

- (66) *deli-k u-ts '-u tfami-s[-ki pudʒi*
 madman-ERG II3.APPL-tell-AOR.13SG pine-DAT-COMPL cow
o-tʃk'om-i]
 PV-eat-AOR
 ‘The madman said to the pine: ‘You’ve eaten the cow’.’ (Ž.105)

Other predicates which use the *ki* strategy are *-dusun-* ‘think (about sth.)’, *-k'itx-* ‘ask’, *-gn-* ‘hear’, *-tkv-* ‘say’ and *-tsk'ed-* ‘see’.

Thirdly, complement clauses may include reported speech marked by one of the clitics *ja*, *ma* or *so* (depending on the person of the superordinate verb), as in (67). This strategy is extended to verbs of thought (68).

- (67) *mi-k g-i-ts '-u [si karmat'e-f nek'na*
 who-ERG II2-APPL-tell-AOR.13SG 2SG mill-GEN door
go-ntsk'-i-ja]
 PV-open-IMP-RS
 ‘Who told you to open the door of the mill?’ (D37.II)

- (68) *tʃkimi feni [i-mt'-asen-ja] g-i-tʃk-i-n-na*
 1SG.GEN for MID-run_away-FUT.13SG-RS II2-APPL-know-TH-13SG-if
 ‘If you think that I’ll run away...’ (lit. ‘if you think about me: ‘She’ll run away’’) (Ž.50)

Fourthly, complement clauses corresponding to content questions do not involve any overt marker. Compare in this respect example (15) with the corresponding indirect question in (69).

- (69) [ma mundes b-yur-are] ko-g-i-tfk-i-n
 1SG when 11-die-FUT.11/2SG PV-II2-APPL-know-TH-13SG
 ‘You know when I’ll die.’ (D67.XXVIII)

Other predicates which use this construction are *oxo-ts’on-* ‘understand’, *go-tfk’ond-* ‘forget’, *-tkv-* ‘say’, *-ts’v-* ‘tell’, *-dzir-* ‘see’ and *do-gur-* ‘learn’.

All the complementation types we have seen so far involve finite verb forms. Laz has a further complementation strategy, which involves the verbal noun (70). Other predicates that use the verbal noun strategy are *-tfk’ond-* ‘forget’ and *-n-* ‘want’.

- (70) [o-p’amit-u-s] ko-gj-o-tfk’-es
 PV-talk-VN-DAT PV-PV-APPL-begin-AOR.13.PL
 ‘They began to talk.’ (K72.136)

6.3. Adverbial clauses

Most adverbial clauses have the structure of postpositional phrases in which the complement of the postposition is a clause rather than an NP. This is illustrated in (71b), which can be compare with (71a), where the postposition takes a nominal complement.

- (71) a. [noya sakis] mende-m-o-on-i-t
 town until PV-II1-TR-take-IMP-1/2PL
 ‘Take me to town!’ (K72.135)
 b. [datfxur-ja do do-b-i-myor-am-t’ā sakis]
 fire-RS and PV-11-MID-shout-TH-SUBJ until
 ge-m-o-ntf’-i-t
 PV-II1-TR-send_down-IMP-1/2PL
 ‘Send me down till I cry ‘fire!’’ (D37.VIII)

As can be seen, adverbial clauses with *sakis* do not include the subordinator *na*. The same structure obtains in adverbial clauses formed with *skule ~ jk’ule* ‘after, while, when, on the point of’ and *doni* ‘since’. Adverbial clauses formed with *feni* ‘in order to, since (cause)’ (ex. 72, to be compared with ex. 54), *steri* ‘as (comparison, manner and time)’ and *k’onari* ‘as much as’ include the subordinator *na*.

- (72) *ma [arka-z-na m-ul-u-nam-pe-na m-a-dzi-a
1SG back-DAT-SUB PV-come-TH-13.PL-PL-SUB II1-POT-see-OPT.13SG
feni] tertsi-z ge-p-x-e guruni-z
for opposite-DAT PV-II1-sit-TH donkey-DAT
'I'm sitting backward on my donkey so that I can see those who come
behind me.'* (own data)

6.4. Conditional clauses

Conditional clauses precede the main clause. They are marked either by *na* or, much less frequently, by *k'on*, which appear at the end of the conditional clause (73).

- (73) [*t/kim k'al ko-mo-xt-a-na*
1SG.GEN with PV-PV-come-OPT-if
si muntxa g-i-no-n do-g-o-gur-are
2SG anything II2-APPL-want-13SG PV-II2-TR-teach-FUT.II/2SG
'If you come with me, I'll teach you anything you like.' (D37.VI)

It results from the discussion of complex clauses that *na* can be considered as a multi-purpose subordinator: it appears in relative, complement and adverbial clauses, where it occurs inside the dependent clause, and in conditional clauses, where it occurs at the end of the dependent clause.

7. Short glossed text

The following text is taken from Lacroix (2009a, text 9). Elements in **bold** are Turkish loans.

- (74) *ar k'otfi **hapisane-se** ama-xt-u-don ama deli jen*
one man prison-ALL PV-go_in-AOR.13SG-EVD but mad be.13SG
'A man went to prison. He is mad.'
- (75) *hek-na i-**tsalif**-am-s **görevli** k'otfi-k **deli-s**
there-SUB MID-work-TH-13SG in_charge man-ERG mad-DAT
u-ts'u-me-s
II3.APPL-tell-TH-13SG*
An employee who works there says to the madman:

- (76) *si ar ts'ana-s zur ts'ana-s otxu ts'ana-s*
 2SG one year-DAT two year-DAT four year-DAT
g-o-tsk'-e do iji he yorma-s o-tsk'-e
 PV-APPL-look-TH and all_the_time DIST.DEM hole-DAT APPL-look-TH
 “You’ve been looking through this hole constantly for one, two, four years.
- (77) *a ndya-s-ti ma me-m-a-/kv-i do*
 one day-DAT-ADD 1SG PV-II1-MID.APPL-let-IMP and
ar ma b-o-tsk'ed-a
 one 1SG II1-APPL-look-OPT
 Let me go there someday, so that I too look at it.”
- (78) *hemu-k-ti zop'on-s-ki o-tsk'ed-i*
 DIST.DEM-ERG-ADD say-i3SG-COMP APPL-look-IMP
 The madman says: “Look”.
- (79) *mola-katʃ-e-epe-s-na o-tsk'-e-n k'otʃi-k-ti*
 PV-shut_up-PART-PL-DAT-SUB APPL-look-TH-i3SG man-ERG-ADD
yorma-s o-tsk'ed-u-/kule mutu va a-dzi-u
 hole-DAT APPL-look-AOR.i3SG-TEMP anything NEG POT-SEE-AOR.i3SG
 When the man in charge of the prisoners looked through the hole, he
 couldn’t see anything.
- (80) *deli-s u-ts'u-me-s görevli-k*
 mad-DAT II3.APPL-tell-TH-i3SG in_charge-ERG
 He said to the madman:
- (81) *hak mutu va ren do*
 here anything NEG be.i3SG and
mu-s o-tsk'-e nak'o ts'ana-s
 what-DAT APPL-look-TH how_many year-DAT
 “There’s nothing here. What have you been looking at for so many years?”
- (82) *ma nak'o ts'ana-s b-o-tsk'ed-i*
 1SG how_many year-DAT II1-APPL-look-AOR
mutu va m-a-dzi-u do
 anything NEG II1-POT-SEE-AOR.i3SG and
 “I looked through the hole so many years and was still not able to see
 anything.
- (83) *si ar defa o-tsk'ed-i do mu dzi-ae-ki*
 2SG one time APPL-look-AOR and what see-FUT.1/2SG-COMP
 You who have looked only once, what could you see?”

Abbreviations

ADD	additive	PERF	perfective
ALL	allative	PL	plural
AOR	aorist	POSS	possessive
APPL	(non-canonical) applicative	POT	potential
CAUS	causative	PROH	prohibitive
COMPL	complementizer	PROX	proximal
DAT	dative	PV	preverb
DEM	demonstrative	RECIPR	reciprocal
DIST	distal	RS	reported speech
ECHO	echoic	SG	singular
EMPH/REFL	emphatic-reflexive	SUB	multi-purpose subordinator
ERG	ergative	SUBJ	subjunctive
EVD	evidential	SUF	suffix lacking a semantic counterpart
FP	future participle	TEMP	temporal suffix or postposition
FUT	future	TH	thematic suffix
GEN	genitive	TR	transitive
IMP	imperative	VN	verbal noun
IMPFT	imperfect	1, 2, 3	1 st , 2 nd , 3 rd person
MID	middle voice	I	Set I person-number features
NEG	negation	II	Set II person-number features
NP	negative participle	x ~ y	x and y are free variants
OPT	optative		
PART	general participle		

References

References to the sources of examples

In the reference of the examples, I indicate the book from which the example was taken and the text number, separated by a dot. Thus, “K72.137” holds for K’art’ozia 1972, text 137. The references of the books are abbreviated as follows:

D37	Dumézil 1937
D67	Dumézil 1967
K72	K’art’ozia 1972
K93	K’art’ozia 1993
Q	Q’ipšidze 1939
Ž	Žyent’i 1938

Other references

- Andrews, Peter A. 1989. *Ethnic Groups in the Republic of Turkey*. Compiled and edited with the assistance of Rüdiger Benninghaus. Wiesbaden: Reichert.
- Dumézil, Georges. 1937. *Contes Lazes*. Paris: Institut d'Ethnologie.
- Dumézil, Georges. 1967. *Documents anatoliens sur les langues et les traditions du Caucase, IV. Récits lazes (dialecte d'Arhavi)*. Paris: Presses Universitaires de France.
- Feurstein, Wolfgang. 1983. Untersuchungen zur materiellen Kultur der Lazen. Freiburg: Albert-Ludwigs Universität, unpublished MA thesis.
- Haig, Geoffrey. 2001. Linguistic diffusion in modern East Anatolia: From top to bottom. In Alexandra Y. Aikhenvald & R.M.W. Dixon (eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, 195–224. Oxford: Oxford University Press.
- Holisky, Dee Ann. 1991. Laz. In Alice C. Harris (ed.), *The Indigenous Languages of the Caucasus*, vol. 3, 395–472. Delmar & New York: Caravan Books.
- K'art'ozia, Guram. 1972. *Lazuri t'ekst'ebi* [Laz texts]. Tbilisi: Mecniereba.
- K'art'ozia, Guram. 1993. *Lazuri t'ekst'ebi II* [Laz texts 2]. Tbilisi: Mecniereba.
- Kutscher, Silvia. 2008. The language of the Laz in Turkey: Contact-induced change or gradual language loss? *Turkic Languages* 12(1). 82–102.
- Lacroix, René. 2009a. Description du dialecte laze d'Arhavi (caucasique du sud, Turquie). Grammaire et textes. Lyon: Université Lyon 2 dissertation. http://theses.univ-lyon2.fr/documents/lyon2/2009/lacroix_r (accessed 12 July 2016).
- Lacroix, René. 2009b. Laz relative clauses in a typological and areal perspective. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds.), *Language Documentation and Linguistic Theory* 2, 205–210. London: SOAS.
- Lacroix, René. 2011. Ditransitive constructions in Laz. *Linguistic Discovery* 9(2). 78–103.
- Lacroix, René. 2012a. Laz middle voice. In Gilles Authier & Katharina Haude (eds.), *Ergativity, Valency and Voice*, 165–198. Berlin & Boston: De Gruyter Mouton.
- Lacroix, René. 2012b. The multi-purpose subordinator *na* in Laz. In Holger Diessel & Volker Gast (eds.), *Clause Linkage in Cross-Linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*, 77–103. Berlin & Boston: De Gruyter Mouton.
- Lacroix, René. 2014. Origin of Sets I-II suffixes in South Caucasian through reanalysis. In Nino Amiridze, Tamar Reseck & Manana Topadze Gäumann (eds.), *Advances in Kartvelian Morphology and Syntax* (Studia Typologica 14), 47–78. Berlin: Akademie Verlag.
- Q'ipšidze, Ioseb. 1939. Č'anuri t'ekst'ebi [Laz texts]. Tbilisi: SSRK' Mecnierebata Ak'ademiis Sakartvelos Pilialis Gamomcemloba.
- Toumarkine, Alexandre. 1995. *Les Lazes en Turquie (XIX^e–XX^e siècles)*. Istanbul: Isis.
- Žyent'i, Sergi. 1938. Č'anuri t'ekst'ebi, arkabuli k'ilok'avi [Laz texts (dialect of Arhavi)]. Tbilisi: SSRK' Mecnierebata Ak'ademiis Sakartvelos Pilialis Gamomcemloba.

6.3. Ossetic

David Erschler

1. Introduction

Ossetic is a cover term for two closely related, but not mutually intelligible Eastern Iranian languages, often called dialects, Iron and Digor¹. Ossetic is spoken in North Ossetia-Alania, an autonomous republic within Russia, and South Ossetia, a breakaway part of Georgia.

In earlier work, Ossetic was considered a member of the Northeastern Iranian subbranch of the Iranian (along with Yaghnobi, Sogdian, and a number of other extinct languages), however, recently doubt has been cast upon the validity of this subgroup. Moreover, Eastern Iranian languages probably do not form a genetic unit either, Sims-Williams (1996).

Systematic research on Ossetic started with Sjögren (1844) and Rosen (1846). Miller (1903) was the first diachronically oriented comprehensive description of Ossetic morphology. Miller (1881) and Miller and Stackelberg (1891) collected and published Ossetic texts with Russian and German translations. The grammar sketch Stackelberg (1886) is based on these texts. Other editions of Ossetic texts with translations into European languages are Christensen (1921) and Munkácsi (1932). Most of later synchronic research on Ossetic was published² in Russian (e.g., Iron Ossetic grammars Axvlediani 1963/1969 and Bagaev 1965/1982; monographic treatments Gabaraev (1977) on word formation; Tekhov (1970) on modality; Kudzoeva (2003) on word order; Tedeev (1989) and Dzodzikova (2009) on verb structure. Digor is addressed in the grammar sketches Isaev (1966); Takazov (2009), and the Digor-Russian dictionary Takazov (2003). Perhaps more accessible to the English-reading audience are Abaev (1964) (although all language data are presented in the Cyrillic-based orthography), a brief general sketch in Thordarson (1989); an overview of phonology by Testen (1997), an overview of word formation Erschler (2015); and a posthumously published

¹ For reasons of space, I do not always provide examples from both languages. When Iron and Digor forms are given simultaneously, they are shown in the following order and separated by a slash: *gədə / tikis* ‘cat’; otherwise examples are marked with (I) and (D) respectively. Data for this chapter were collected in the course of my fieldwork in North Ossetia in 2007–2013. I am grateful to all my consultants for their generous helpfulness. I have worked particularly much with Arbilana Abaeva, Tsara Dzhanaev, and Elizaveta Kochieva (Iron); and with Sveta Gatieva, Marina Khamitsaeva, Khasan Maliev†, Chermen Takazov, Fedar Takazov, and Vera Takazova (Digor).

² Unpublished dissertations are not listed here, as they are virtually inaccessible even in Russia.

collection of sketches Thordarson (2009). Furthermore, Haspelmath (1997) has an overview of indefinite pronouns in Iron Ossetic. Some of the recently published papers in English are Belyaev and Vydrin (2011); Erschler (2009, 2012a, 2012b); Erschler and Volk (2011); and Tomelleri (2009).

Ossetians are descendants of medieval Alans and, through the Alans, of Sarmatians. Their settlement in the piedmont of the Caucasus dates back to the first centuries of the Common Era. Before that, nomadic ancestors of Ossetians wandered in Eurasian steppes between the Danube and the Altay mountains. In early middle ages, the Christian Alanic kingdom controlled much of the north-western Caucasus. It was destroyed in the 13th–14th c. by Mongols and Timur, Abaev and Bailey (1985). Following the kingdom's destruction, the ancestors of Ossetians only survived in high mountain valleys, and, until the early 18th c., data about them is extremely scarce. By the early 18th c., their territory roughly corresponded to the highlands of the modern South and North Ossetia, Volkova (1974). A subgroup of Alans, known as Iassi, migrated to Hungary in the 13th or 14th c. Although they lost the language at some point, their descendants remained a separate ethnic group within Hungary at least until the early 20th c.

In the recent past, Ossetians lived in high mountain valleys and mainly practiced animal husbandry. However, in the course of the last two centuries, the majority of the population moved to the plains and took up agriculture. Currently, their lifestyle is considerably urbanized.

The earliest written attestations of Alanic/early Ossetic are an Alanic tombstone inscription from the tenth to twelfth century, a few lines in Alanic in a twelfth century Byzantine manuscript *Theogony* by Ioannes Tzetzes, a few fourteenth or fifteenth century glosses in a Greek manuscript (Alemany 2000; Kambolov 2006: 166–207) and an Iassic word list recorded in Hungary in the 15th c., Németh (1959). The first “real” texts date back to the late eighteenth to early nineteenth century (translations from Church Slavonic and Georgian).

Ossetians form the majority of the population in both North and South Ossetia. Most of them are Iron speakers (modulo the language loss of the last half-century), but no precise figures are available. According to the 2010 census, there are 528,000 Ossetians (Iron and Digor) in Russia. Moreover, an unknown number of Iron speakers live in South Ossetia, probably between 30,000 and 70,000. Also, a few thousands Ossetians (Iron and Digor) live in Turkey. Digor Ossetians live mostly in Digor and Iraf districts of North Ossetia. The total population of these districts according to the 2010 census is about 35,000; this figure can be interpreted as a rough estimate of the number of Digor speakers.

Digor has two major dialects, one spoken in the town of Digora and another in the villages of Chikola, Sheker, Dzagepbarz, Khazhnidon, Toldzgun, and Lesken (see Figure 1). They differ in phonetics and, probably, to some extent in morphosyntax as well. It is the Chikola-Lesken dialect of Digor that is represented in this chapter. Moreover, there was a mixed Iron-Digor dialect spoken in the Wallagkom

valley (Isaev 1966), which is now virtually extinct. Another mixed Iron-Digor dialect emerged in the village of Ursdon-Sindziqaw. No systematic study of Digor dialects has been undertaken so far.

Iron has about half a dozen dialects (Bekoev 1985). A major South Ossetian dialect, Kudar, is sometimes perceived as the third Ossetic language. While phonetic differences between Iron dialects are fairly well described, morphosyntactic differences have not been studied systematically. All dialects use the same orthography, which does not reflect the phonetic differences.

Nominally, the majority of North Ossetians and virtually all South Ossetians are Orthodox Christian, and, in North Ossetia, there exists a sizeable Moslem minority. In fact, the traditional syncretic religion is still widely practiced.

In North Ossetia, no monolingual speakers remain, and at least 40% of Ossetians are more fluent in Russian than in Ossetic according to Kambolov (2007). Kambolov's data are based on respondents' self-evaluation, and the actual percentage of fluent speakers may be significantly lower. In North Ossetia, Ossetic is now effectively restricted to the private sphere. Street signs, price tags in shops or announcements in public transportation are all in Russian, which is also the language of local authorities and courts.

The language situation in South Ossetia is somewhat better in this respect. While the main language of education is Russian as well, in some rural schools the language of education is at least partly in Iron for first several years. However, it then completely switches to Russian and Iron is only taught only as a subject. The situation with Digor is even worse.

In rural areas, Ossetic is still acquired by children. However, about 50% of all Ossetians live in Vladikavkaz, the Russian speaking capital of North Ossetia, and children raised there usually grow up semispeakers at best. Furthermore, a high level of unemployment in rural areas causes migration to Vladikavkaz and to inner regions of Russia. All this makes the long term survival of Ossetic questionable.

Both Iron and Digor are written languages, books and periodicals are published, and the Vladikavkaz TV broadcasts short programs in Iron and, since recently, in Digor. Between the mid-1950s and the early 1990s, publishing in Digor was strongly discouraged.

The modern Ossetic orthography is Cyrillic based, with only one additional letter *Æ/æ* for the sound *v*; the rest of specifically Ossetic sounds are expressed by digraphs. Ways of transcribing/transliterating Ossetic vary widely across the literature.

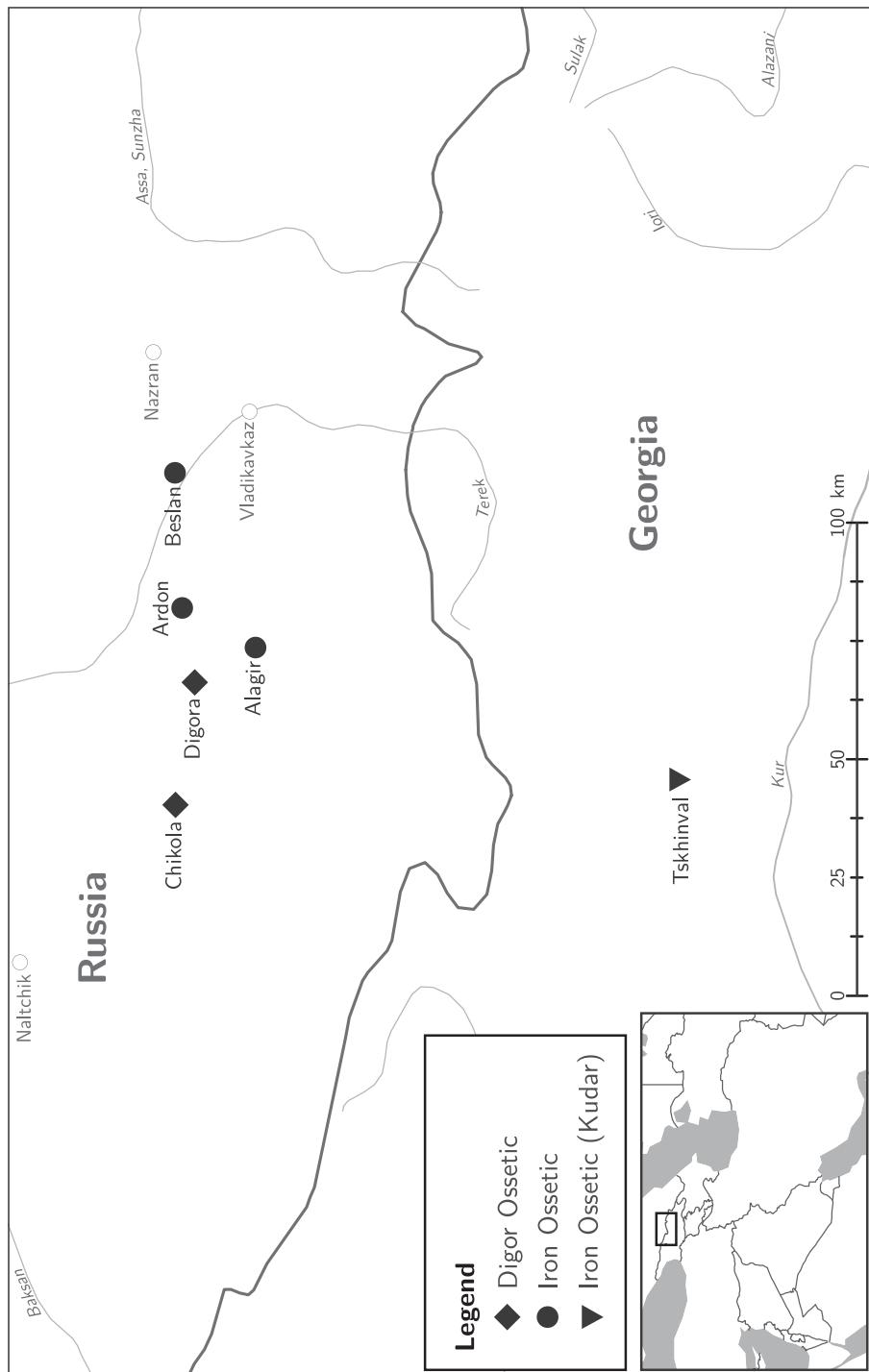


Figure 1: Diger and Iron Ossetic locations mentioned in this chapter

2. Phonology

2.1. Segmental phonology

Ossetic phonology is described in Testen (1997). A recent instrumental study of Iron phonetics is Dzakhova (2009). Typically for the area, Ossetic has a series of ejective consonants *p'*, *t'*, *c'*, *č'* and *k'*, which mostly appear in loans or onomatopoeic words, but also developed in some words of Indo-European origin, e. g. *št'ala/st'alu* ‘star’, *lašk'/lisk'v* ‘nit’. This series of consonants is obviously a borrowing from autochthonous languages of the Caucasus.

Vowel inventory of Digor Ossetic

<i>i</i>	<i>u</i>
<i>e</i>	<i>o</i>
<i>ɛ</i>	
<i>a</i>	

Vowel inventory of Iron Ossetic

<i>i</i>	<i>u</i>
	<i>ə</i>
<i>e</i>	<i>o</i>
<i>ɛ</i>	
<i>a</i>	

There is no phonemic contrast in length. However, /ɛ/, /i/, /u/, and /ə/ are acoustically shorter than the rest of the vowels, Sokolova (1953).

Table 1: Consonant inventory

	labial		coronal			velar	uvular	glottal
	bilabial	labiodental	dental	alveolar	postalveolar			
stop	b		d(=g)			g, g ^w	q q ^w	?
	p(=p ^h)		t(=t ^h)			k(=k ^h), k ^w		
	p'		t'			k' k' ^w		
fricative		f v		s z	ž (=ʒ) š (=ʃ)		ʂ ʂ ^w	h
							χ χ ^w	
affricate			dz		dʒ č(=tʃ)			
			c(=ts)		č'(=tʃ')			
			c'(=ts')					
nasal	m			n				
lateral				l				
trill				r				
glide	w				j			

In this chapter, I will not be using the IPA symbols in parentheses in Table 1.

In standard Iron, /c/ only survives in certain clusters. Glottal consonants are marginal; the glottal stop /ʔ/ is automatically epenthesized before word-initial /a/, /ə/, and /ə/, and sometimes between vowels. The glottal fricative /h/ is present in very few words: *ho* ‘yes’, *henər/henur* ‘now’, *maha* ‘I don’t know’, *haj-haj* ‘certainly, of course’ and some others.

All consonants are phonetically palatalized before front vowels, in Digor the palatalization leads to changes in the primary articulation for the sibilants and affricates: /s/ becomes /š/, /z/ becomes /ž/, and /c/ becomes /č/. If a front vowel follows a cluster, then all the consonants in the cluster are palatalized: D. *kust* ‘work’ > *kušt-itə* work-PL. This happens also when an enclitic with a front vowel is attached to a word: *k’os=mi* > *k’oš=mi* cup=ABL.1SG. In Iron, the velar stops, /k/ and /g/, become /č/ and /dž/, respectively, before the oblique case marker -ə.

2.2. Phonological processes

Consonant assimilation is only observed when the allative marker *-mə* is attached to an *n*-final stem, in which case the /n/ becomes /m/. Ossetic disallows hiatus, and a variety of epenthetic consonants (usually glides or the glottal stop) are used to break it up. Word finally, voiced consonants undergo devoicing. Vowel reduction as such has not been attested, but, in Iron, /ə/ in some positions can be inaudible in fluent speech.

2.3. Phonotactics

Minimal lexical words have CVC structure. The 2SG pronoun *də/du* and some wh-words are CV (e. g. D. *ka* ‘who.NOM’; *či* ‘what.NOM/ACC’). Wh-words can function as non-clitics in fragment questions, but normally they are procliticized to a verb. The rest of CV and VC entities are either clitics or bound morphemes.

There is no reliable evidence that the notion of syllable as such is relevant for Ossetic grammar. Bailey (1950) claimed that syllable weight is relevant for Digor stress, but his stress data appear to be wrong. Initial clusters contain no more than two consonants *št’alə/st’alu* ‘star’. In some Iron dialects, initial clusters are resolved by the prosthetic ə: *əšt’alə*. In Digor, the prosthetic vowel is ε. Final clusters³ of the form *nC*, *lC*, and *rC*, where C is a non-labial stop, are relatively common. Other types of cluster are very rare except for a large number of clusters with the final *d/t* that appear in past participles. Historically, the *-d/-t* was the past participle exponent. In these participles, any consonant except the stops and the affricates may precede the *d/t*. Three-consonant clusters are attested only in such participles; the penultimate consonant then is always *s*: Digor *vχst* ‘shot’.

³ Due to the limitations of space, I only describe Digor clusters here.

2.4. Stress and prosody

Both Digor and Iron lack lexical or grammatical tone. The acoustic correlates of stress are unknown. Dzaxova (2010) states that in Iron, the acoustic correlate is a higher pitch on the respective syllable, but this does not hold for all words of her sample (and the words were recorded without a carrier sentence). For Digor, no conjectures have been advanced so far and a pilot study carried out by the author has failed to detect lexical stress. There are no modern studies of prosody in Ossetic, although some observations can be found in Abaev (1949).

3. Morphology

For an overview of word formation in Ossetic, see Erschler (2016) and the references therein. Open word classes are verbs and nouns.

3.1. Nominal morphology

Adjectives and adverbs are morphologically indistinguishable from nouns. Other classes of nominals include numerals, full pronouns, enclitic pronouns, reflexives (also used as intensifiers), reciprocals, wh-words, various types of indefinites, and quantifiers.

Nouns are inflected for *number* and *case*; the markers are suffixed to the noun, and, noun phrases being strictly head-final, to the right edge of noun phrases. There is no *grammatical gender*, even in pronouns. *Animacy* is not marked overtly, but plays a role in differential object marking. *Possessedness* is marked by a prefix/proclitic, which attaches to the left edge of a noun phrase and distinguishes person and number of the possessor.

- (1) *jə= asə štər wurš bəχ-t-ən* (I)
 POSS.3SG= this big white horse-PL-DAT
 ‘for these big white horses of his/hers’

The plural marker attaches immediately to the right of the stem. It may trigger a vowel alternation, which is not completely predictable, and, furthermore, if the noun ends in a sonorant, the *-t-* of the plural marker may be geminated: (D) *bon* ‘day’ vs. *ben-tte* ‘days⁴’. If a word ends in a cluster, the allomorph of the plural is *-ətə/-itə*, with a possible vowel alternation in the stem: *vəχst/vəχst* ‘shot’ → *vəχst-ətə/vəχst-itə*. Plural marking is possible for wh-phrases and some indefinites as well: D. *či-te* what-PL; *ka-der-te* who-IDF-PL lit. ‘someones’. Semantic conditions under which it arises are unclear.

⁴ The regular form *bon-tə* exists as well.

Personal names may carry the associative plural marker *-i-/e-*: Ir. *Aslan-i-te* Aslan-ASS-PL ‘Aslan and others with him’.

Case is usually marked on the right edge of the word form. The only exceptions are certain indefinites, in which the case marker either precedes the indefinite suffix (e. g. the Digor superessive form of *či-der* what-IDF ‘something’ is *cebel-der* what.SUP-IDF) or, for plural marked indefinites, occurs several times in the word form (e. g. the Digor ablative form of *ka-der-te* who-IDF-PL ‘someone-PL’ is *kemej-der-t-i / kemej-der-t-ej* who.ABL-IDF-PL-OBL/ABL.) For nouns, numerals, and quantifiers the case marking is nearly agglutinative, only minor morphonological alternations occur. Prenominal stems show suppletion, see paradigms below. In Digor, numerals require dedicated numeral declension suffixes inserted between the stem and the regular case marker, for example *duw-e-j* ‘two-NUM-OBL’. Remarkably, numeral phrases join the numeral class: *bəχ-i* horse-OBL, but *duwb bəχ-e-j* two horse-NUM-OBL.

Case markers differ from postpositions in the following respects: a) case markers attach to bare noun stems, whereas postpositions require case marking on nouns (usually the oblique); b) case-marked nouns, unlike postposition complements, may bind clause-mate reflexives and reciprocals, and control depictives, Erschler (2012b).

Ossetic does not allow Suffixaufnahme, i. e. stacking of several case suffixes on a single noun phrase: *NP-CASE₁-CASE₂. Case affixes may be suspended, i. e. one case affix may mark two conjoined noun phrases, Erschler (2012b).

- (2) *kʷəz emə gedə-mə* (I)
dog and cat-ALL
'to a cat and a dog'

When a pronoun is coordinated with a noun, suspended affixation works slightly differently: the first conjunct receives the oblique case, while the second takes the case relevant to the entire phrase (see paradigms below), as in (3). When two pronouns are coordinated, suspended affixation is ungrammatical.

- (3) *dəw/*?də emə aslan-en* (I)
you.OBL/you.NOM and Aslan-DAT
'to you and to Aslan'

The Ossetic case system consists of the nominative; accusative/genitive/inessive; dative; ablative; allative; superessive; equative. Iron also has the comitative. The accusative, genitive and inessive coincide for lexical nouns. The accusative and the genitive only differ for clitics: (second position) enclitics correspond to direct objects, whereas proclitics to noun phrases correspond to possessors (Erschler 2009). The inessive differs from the genitive and accusative for enclitics, and, in Digor, for numerals and numeral phrases. In this chapter, I gloss the syncretic marker of the accusative, genitive, and inessive as the oblique.

Table 2: Noun inflection (Digor)

Consonant-final		v-final	Other vowel-final
	Plurals	Singular	
Nom	<i>bəχ</i> ‘horse’	<i>bəχ-tə</i> ‘horse’	<i>kizgə</i> ‘girl’
Obl	<i>bəχ-i</i>	<i>bəχ-t-i</i>	<i>kižg-i</i>
Dat	<i>bəχ-en</i>	<i>bəχ-t-en</i>	<i>kizg-en</i>
Abl	<i>bəχ-ej</i>	<i>bəχ-t-ej</i>	<i>kizg-ej</i>
All	<i>bəχ-me</i>	<i>bəχ-tə-me</i>	<i>kizgə-me</i>
Sup	<i>bəχ-bel</i>	<i>bəχ-tə-bel</i>	<i>kizgə-bel</i>
Equ	<i>bəχ-aw</i>	<i>bəχ-t-aw</i>	<i>kizg-aw</i>

Table 3: Inflection of personal pronouns (Digor)

	1		2	
	Sg	Pl	Sg	Pl
Nom	<i>ez</i>	<i>maχ</i>	<i>du</i>	<i>sumax</i>
Obl	<i>men</i>	<i>maχ</i>	<i>dəw</i>	<i>sumax</i>
Dat	<i>men-en</i>	<i>maχ-en</i>	<i>dəw-en</i>	<i>sumax-en</i>
Abl	<i>men-ej</i>	<i>maχ-ej</i>	<i>dəw-ej</i>	<i>sumax-ej</i>
Iness	⁵	—	—	—
All	<i>menme</i>	<i>maχ-me</i>	<i>dəw-me</i>	<i>sumax-me</i>
Sup	<i>men-bel</i>	<i>maχ-bel</i>	<i>dəw-bel</i>	<i>sumax-bel</i>
Equ	<i>men-aw</i>	<i>maχ-aw</i>	<i>dəw-aw</i>	<i>sumax-aw</i>

Table 4: Inflection of deictics.⁶ (Digor)

Nom	<i>a/ajə</i>	<i>a-ɪə</i>	<i>je</i>	<i>je-tə</i>
Obl	<i>aj</i>	<i>an-i</i>	<i>woj</i>	<i>won-i</i>
Dat	<i>a-m-en</i>	<i>an-en</i>	<i>wo-m-en</i>	<i>won-en</i>
Abl	<i>a-m-ej</i>	<i>an-ej</i>	<i>wo-m-ej</i>	<i>won-ej</i>
Iness	<i>a-m-i</i>	<i>an-em-i</i>	<i>wo-m-i</i>	<i>won-e-m-i</i>
All	<i>a-me</i>	<i>an-emə</i>	<i>wo-me</i>	<i>won-e-me</i>
Sup	<i>a-bel</i>	<i>an-ebel</i>	<i>wo-bel</i>	<i>won-e-bel</i>
Equ	<i>a-j-aw</i>	<i>ani-j-aw</i>	<i>wo-j-aw</i>	<i>woni-j-aw</i>

⁵ For the 1st and the 2nd persons, localization “in” can only be expressed with a postposition.

⁶ The distal deictics *je/jetə* also serve as the 3rd person pronouns.

Table 5: Inflection of enclitics (Digor)

	1		2		3	
	Sg	Pl	Sg	Pl	Sg	Pl
Nom	–	–	–	–	–	–
Acc	<i>me</i>	<i>ne</i>	<i>de</i>	<i>we</i>	<i>ej</i>	<i>se</i>
Gen	–	–	–	–	–	–
Dat	<i>min</i>	<i>nin</i>	<i>din</i>	<i>win</i>	<i>in</i>	<i>šin</i>
Abl	<i>mi</i>	<i>ni</i>	<i>di</i>	<i>wi</i>	<i>ši</i>	<i>ši</i>
Iness	–	–	–	–	<i>ši</i>	<i>ši</i>
All	<i>meme</i>	<i>neme</i>	<i>deme</i>	<i>weme</i>	<i>ime</i>	<i>sema</i>
Sup	<i>mebel</i>	<i>nebel</i>	<i>debel</i>	<i>webel</i>	<i>ibel</i>	<i>sebel</i>
Equ	–	–	–	–	–	–

Table 6: Inflection of numerals and numeral phrases (Digor⁷)

Nom	<i>duw-e</i> '2' <i>insej</i> '20'	<i>duwe</i>	<i>bon-i</i> day-OBL
Obl	<i>duw-e-j</i> <i>insej-e-j</i>	<i>duwe</i>	<i>bon-e-j</i> day-NUM-OBL
Abl	<i>duw-em-ej</i> <i>insej-em-ej</i>	<i>duwe</i>	<i>bon-em-ej</i> day=NUM-ABL
Dat	<i>duw-em-en</i> <i>insej-em-en</i>	# <i>duwe</i>	<i>bon-em-en</i> ⁸ day-NUM-DAT
Iness	<i>duw-em-i</i> <i>insejem-i</i>	# <i>duwe</i>	<i>bon-em-i</i> day-NUM-INES
All	<i>duw-e-me</i> <i>insej-em-e</i>	<i>duwe</i>	<i>bon-e-me</i> day-NUM-ALL
Sup	<i>duw-e-bel</i> <i>insej-e-bel</i>	<i>duwe</i>	<i>bon-e-bel</i> day-NUM-SUP
Equ	# <i>duw-e-j-aw</i> # <i>insej-aw</i>	# <i>duwe</i>	<i>bon-e-jaw</i> day-NUM-EQU

⁷ In Iron, numerals are declined identically to nouns. In numeral phrases in the nominative, the oblique marking surfaces on the nouns: *ertə qaž-a* three goose-OBL 'three geese'. In other cases, the oblique is replaced by the respective case marker: *ertə qaž-en* three goose-DAT.

⁸ The # sign marks morphologically possible but pragmatically improbable forms.

3.2. Verbal morphology

Ossetic exhibits a robust difference between finite and non-finite forms: only finite forms agree with their subjects in *person* and *number*, and show *tense* and *mood* distinction. The *aspect*, on the other hand, is defined for non-finite forms as well. For an overview of non-finite forms, see Erschler (2015).

Compared to other languages of the area, Ossetic only shows a rather small array of finite forms: the present, past and future of the indicative, the imperative, and the three tenses of the subjunctive. Admittedly, for the non-indicative forms, the terms past/present/future subjunctive are convenient labels rather than terms meaningfully reflecting the semantics of these forms. The past subjunctive, besides its use in conditional and counterfactual clauses, can express estimate or habituality in the past. The latter is also the most common use of the present subjunctive in Digor. Unusually for the area, Ossetic lacks grammatical evidentiality.

Except for a handful of roots, preverbless verbs are imperfective, and the perfective aspect is marked by preverbs. In Iron, perfective verbs can be imperfectivized or endowed with the meaning ‘had almost done x’ by inserting the conative affix *-sej-* between the preverb and the root. In Digor written texts, analogous forms with *-cej-* are sometimes encountered, but they have been rejected as Iron calques by the majority – in fact, all but one – of the native speakers I have worked with.

- (4) a. *južaman me=sigaret-t-ə k'opp fe-sej-fewud kodta* (I)
 once my=cigarette-PL-OBL pack PRV-CON-finish
do.PST.3SG
 ‘Once, I was running out of cigarettes.’ (Lit. ‘Once, my pack of cigarettes was ending.’)
 (Dzasoxt 2007)
- b. *o xedegaj kʷəd=mə fe-sej-rox kodta* (I)
 yes by_the_way how=ABL.1SG PRV-CON-forgetting
do.PST.3SG
 ‘Yes, by the way, I had almost forgotten.’
 (Dzasoxt 2007)

Besides these basic forms, there exists a plethora of analytic modal constructions, see some examples in Belyaev and Vydrin (2011) and in Erschler (2014). A periphrastic form (the converb plus a finite form of the verb ‘to do’, which then carries all TAM marking) is used when the lexical verb has to be focused or topicalized:

- (5) *win-gr=bə=vj ma fe-kken-ed* (D)
 see-CVB=CTR=ACC.3SG NEG.IMP PRV-do-IMP.3SG
 ‘Let him/her never see it.’
 (Gurdzibety 2006)

A finite verb only agrees with the nominative argument irrespective of the semantic role of the latter. The agreement is in person and number; collective nouns may trigger plural agreement, whereas plural noun phrases may agree in the singular. Details of the phenomenon have not been systematically studied.

Only about 300 verbs are simplex (i. e. attach the agreement markers directly to the stem), the rest are a combination of a nominal part with a light verb, usually ‘to do’ or ‘to be’: Ir. *axʷər kən-* learning do ‘to study’. Simplex verbs have two stems, the past and the present, the relationship between which is irregular (e. g. Ir. *səw-/səd-* ‘go’ vs. *ləw-/ləwəd-* ‘stand’), and they both have to be listed in the lexicon.

For simplex finite verbs, the maximal structure is preverb-conative suffix-stem-tense.mood.agreement, of which the preverb is optional, and the conative suffix may be inserted only in the presence of a preverb. For complex verbs, preverbs, and, possibly, the conative suffix attach to the nominal part, whereas tense-mood-agreement suffixes are carried by the light verb.

If a verb occurs sentence-initially, then, in Iron, the preverb with the conative suffix *-stj-* can be separated from the root by a pronominal clitic, as in (6a). In Digor, this is possible for bare preverbs, (6b). In (6b), *ra-jevṣud un* ‘to pass (about time)’ is a complex verb, whose nominal part carries the prefix *ra-*.

- (6) a. *fə-səj=šəm* *χəccə* *kod-ton* (I)
 PRV-CON=ALL.3PL near do.PST-PST.1SG
 ‘I was approaching them.’
 (Bedzhyzaty 1995: 55)

b. *ra=məbel* *jevəud* *vj* *fezzəg* (D)
 PRV=SUP.1SG passed be.PRS.3SG autumn
 ‘I spent the fall.’ (Lit. ‘The fall passed on me.’)
 (Aghuzarti 2008)

Verbs fall into two conjugation classes, the difference between which only shows in past indicative forms: Ir. *kod-toj* do.PST.3PL vs. *səd-əšta* go.PST.3PL. They are often called transitive and intransitive in the literature, presumably because the majority of verbs in each class are transitive and intransitive respectively. However, class assignment is lexically determined, and there are exceptions. For example, in Iron, the verbs *žar-* ‘to sing’ and *keš-* ‘to read’ are morphologically intransitive: *žarəd-əšta* sing.PST-PST.3PL ‘they sang’, *kašt-əšta* read.PST-PST.3PL ‘they read’, whereas *rej-* ‘to bark’ and *felzvəd-* ‘to snow heavily, especially during a blizzard’ are transitive: *rejd-toj* bark.PST-PST.3PL ‘they barked’; *mit felzvək-ta* snow snow.PST-PST.3SG ‘It snowed heavily.’

In Iron, iterativity/habituality in the past is usually marked by the second position clitic *-ju* (in some environments pronounced *-iw*), whereas Digor mostly uses the past and present subjunctive forms for this. The habituality in the present is expressed by preverbs, (7). In the following example, the preverbs and the present

indicative morphology are highlighted in boldface. Concerning the relative clause structure, see §5.4 below.

- (7) [nur=dər=ma ješke džubandi-tə ke zverde-me
 now=EMP=still somebody.OBL word-PL who.OBL heart-ALL
 ne=fə-cce-nce] /je wote ba-ken-uj (...) (D)
 NEG=PRV-go-PRS.3PL he.NOM so PRV-do-PRS.3SG
 ‘Even now, (whenever) somebody’s words would not please one, s/he
 would say so (...).’
 (Sabajti 2010)

The reflexive, reciprocal, and applicative are not marked on the verb: the reflexive and reciprocal use dedicated pronouns, whereas the argument introduced by the applicative is marked by the dative. Unlike in Slavic or Northwestern Caucasian, Ossetic preverbs do not introduce arguments. The causative is expressed analytically, as the infinitive of the lexical verb plus the auxiliary ‘to do’: Ir. *ba-mbar-aŋ* *kodta* PRV-understand-INF do.PST.3SG ‘s/he explained’. The passive is fairly rare; it is expressed as the past participle or the converb plus the appropriate finite form of the copula or of the verb ‘to go’.

- (8) a. sovetton xecawadə terk-i bəstə-bal sast
 Soviet authority Terek-OBL region-SUP break.PRT.PST
 ku (v)r-cud-ej
 when PRV-go.PST-PST.3SG
 ‘When the Soviets had been defeated (lit. broken) across Terek
 region...’⁹
- b. sk'ola zumeg-i wadz-gr er-cud-ej (D)
 school winter-OBL leave-CVB PRV-go.PST-PST.3SG
 ‘The school had to be abandoned in winter.’

Below, Digor verb paradigms are illustrated.

The Paradigm of the Copula

Table 7: Indicative

	Present		Past		Future	
1	<i>dən</i>	<i>an</i>	<i>adtən</i>	<i>adtən</i>	<i>wodžan(en)</i>	<i>wodž(in)an</i>
2	<i>də</i>	<i>ajtə</i>	<i>adtə</i>	<i>adtajtə</i>	<i>wodžene</i>	<i>wodž(in)ajtə</i>
3	<i>vj / jes</i> ¹⁰	<i>enče</i>	<i>adtej</i>	<i>adtənce</i>	<i>wodže(ne)j</i>	<i>wodž(en)enče</i>

⁹ (http://allingvo.ru/DIGOR/bagharati_sozur.htm, accessed 08 July 2017).

¹⁰ In PRS.3SG, the final -s may be dropped, and the existential copula is often pronounced *je*.

Table 8: Imperative

	SG	PL
2	<i>wo</i>	<i>wote</i>
3	<i>wəd</i>	<i>wente</i>

Table 9: Subjunctive

	Present		Past		Future	
	SG	PL	SG	PL	SG	PL
1	<i>wine</i> (<i>wajine</i>) ¹¹	<i>wijane</i> (<i>wajane</i>)	<i>adtajne</i>	<i>adtajanə</i>	<i>won</i>	<i>wen</i>
2	<i>wise</i> (<i>wajse</i>)	<i>wiajte</i> (<i>wajajte</i>)	<i>adtajse</i>	<i>adtajajte</i>	<i>waj</i>	<i>wajte</i>
3	<i>wide</i> (<i>wajde</i>)	<i>wiwonce</i> (<i>wajwonce</i>)	<i>adtajde</i>	<i>adtajiwonce</i>	<i>wa</i>	<i>wonce</i>

Paradigms of simplex verbs, Digor. (Unless indicated otherwise, illustrated for ‘to do’)

Table 10: Indicative

	Present, PRS stem		Future, PRS stem	
	both conjugations		both conjugations	
1	<i>ken-un</i>	<i>ken-en</i>	<i>ken-dz(en)en</i> ¹²	<i>ken-dʒ(in)an</i>
2	<i>ken-i(s)</i> ¹³	<i>ken-(e)</i> ¹⁴ <i>te</i>	<i>ken-dz(en)e</i>	<i>ken-dʒ(in)ajte</i>
3	<i>ken-uj</i>	<i>ken-uncə</i>	<i>ken-dz(en)ej</i>	<i>ken-dʒ(en)ence</i>

Past, PST stem			
“Transitive” conjugation, ‘to do’		“Intransitive” conjugation, ‘to look’	
<i>kod-ton</i>	<i>kod-tan</i>	<i>kas-ten</i>	<i>kast-an</i>
<i>kod-taj</i>	<i>kod-tajte</i>	<i>kas-te</i>	<i>kast-ajte</i>
<i>kod-ta</i>	<i>kod-tonce</i>	<i>kas-tej</i>	<i>kast-ənce</i>

¹¹ *wi-* and *wa-* forms of the present subjunctive appear to be completely equivalent, but *wi-* forms are much more frequent. Some younger speakers do not recognize the variant with *wa-*.

¹² The “long” and “short” forms of the future suffix are completely equivalent, in careful speech and writing it is mostly the longer form that is used.

¹³ The final *-s* of PRS.2SG is usually dropped in colloquial speech.

¹⁴ *-e-* is optional here only for this verb.

Table 11: Imperative (present stem)

	Sg	Pl
2	<i>ken-e</i>	<i>ken-(e)te</i>
3	<i>ken-ed</i>	<i>ken-ente</i>

With some verbs, the imperative is irregular:

cəw-un ‘go’ IMP.2SG *co/*cəwe*; IMP.2PL *co-te/cəwete*;

kew-un ‘to weep’ IMP.2SG *ko/kewə* IMP.2PL *ko-te /kewə-te* The respective 3rd person forms are regular.

Table 12: Subjunctive (same for both conjugations)

	Present (PRS stem)		Past (PST stem)		Future (PRS stem)	
1	<i>ken-inə</i>	<i>ken-ijanə</i>	<i>kod-tajne</i>	<i>kod-tajane</i>	<i>ken-on</i>	<i>ken-an</i>
2	<i>ken-isə</i>	<i>ken-ijajte</i>	<i>kod-tajse</i>	<i>kod-tajajte</i>	<i>ken-aj</i>	<i>ken-ajte</i>
3	<i>ken-idə</i>	<i>ken-iwonce</i>	<i>kod-tajde</i>	<i>kod-tajiwonce</i>	<i>ken-a</i>	<i>ken-once</i>

4. Basic clause structure

4.1. Noun phrase and adpositions

Noun phrases are strictly head final and unsplittable, even by clitics. There is no overt agreement within noun phrases.

- (9) *mə=wəsə dəwwər χorž ləmen-an* (I)
 POSS.1SG=that two good friend-DAT
 ‘for these two good friends of mine’

Numerals precede noun phrases they quantify, which stand in the singular.

There is only one preposition, *ene* ‘without’. The majority of postpositions are transparently related to lexical nouns, e.g. *šer/sər* ‘head (n.)’; on (postp.). With verbs of movement, postpositions necessarily carry plural marking:

- (10) *χed-i sər-* (t)-i ba-wad-ten* (D)
 bridge-OBL on-PL-OBL PRV-pass-PST.1SG
 ‘I passed over a bridge.’

Postpositions can be case-marked and assign case, most commonly the oblique, to their complement noun phrases. For oblique-assigning postpositions,¹⁵ the complement can be realized as a possessive proclitic: Ir. *mə=šər-mə* POSS.1SG=on-ALL

¹⁵ Except *təχχəj/tuχχəj* ‘because of, about’, which only allows non-clitic complements.

'above me'. For postpositions assigning other cases, their complements cannot be replaced by clitics.

4.2. Verb phrase

Verb phrase, as a surface constituent, does not exist (although there might be arguments for positing it on a more abstract level).

4.3. Constituent order

Non-finite clauses are rigidly verb final; the word order in finite clauses is in principle free, but SOV largely prevails, modulo the usual disclaimer that clauses where all arguments are expressed by full noun phrases are relatively infrequent. Ossetic shows pro-drop, whereas discourse-given non-subject noun phrases are usually expressed by clitics.

4.4. Argument marking

By default, Ossetic is strictly nominative-accusative. There are no instances of TAM-based alignment splits. Very few verbs have different argument marking patterns. These verbs either agree with the stimulus or show the default 3sg agreement. They include: *fend-* ‘to want’; *qew-/kew-* ‘to need’, and, in Iron, *wərn-* ‘to believe’, which mark their experiencers with the oblique. The verb *enteš-/entes-* ‘to succeed’ marks the experiencer with the dative. In Digor, the verb *enkez-*, lit. ‘to leaven’ is used with the default 3SG morphology and a dative “subject” in the meaning “to be permitted to X”. Furthermore, in one of the diatheses, the complex verb *PRV-rox / PRV-ironx COP* ‘to forget’ marks the experiencer with the ablative:

- (11) *fe-ronχ=mi* *vj*
 PRV-forgetting=ABL.1SG be.PRS.3SG
 ‘I have forgotten it.’

Ossetic shows differential object marking. Referential animate direct objects are obligatorily marked with the oblique, (12a–b), while inanimate direct objects normally remain unmarked, (12e) but the details are as yet poorly understood (Thordarson 2009: 135–137). As the sentences in (12c–d) show, if the animate object is non-specific or non-referential, the object stays in the nominative, whereas when both the subject and the object are inanimate, the latter gets the oblique marking.

- (12) a. Animate definite direct object

leppu-jə vddər-mə ra-šajd-toj (I)

boy-OBL outside-ALL PRV-deceive.PST-PST.3SG

‘They teased the boy outside (by deceit).’

(Qajtyqty 1998)

- b. Animate referential direct object

k'wəri raždər wəm bərəždʒ-ə gal-ə birek-te (I)

week earlier there Barazg-OBL bull-OBL wolf-PL

erkevš-toj

slaughter.PST-PST.3PL

‘A week ago, wolves killed Barazg’s bull there.’

(Qajtyqty 1998)

- c. Animate non-referential direct object

kušərttag-ən=iw š-fəχ-toj gal (I)

slaughtered-DAT=HAB PRV-cook.PST-PST.3SG bull

‘Usually, they cooked a bull (as the animal for a feast).’

(Qajtyqty 1998)

- d. Inanimate subject, inanimate object

wad a-šaš-ta gədə bəlaš-(ə)* (I)

storm PRV-break-PST.3SG cat tree-OBL

‘The/a storm broke the/a poplar.’

(Bedzhyzaty 1995)

- e. Animate subject, inanimate direct object

leppu a-šaš-ta gədə bəlaš-(?ə) (I)

boy PRV-break-PST.3SG cat tree-OBL

‘The/a boy broke the/a poplar.’

4.5. Clitics

Enclitic pronouns (together with some other enclitics) form a cluster that obligatorily occupies (an appropriately defined) second position in Iron.

- (13) *nə=tə=jən=vj* (I)

NEG=again=DAT.3SG=ACC.3SG give.PST.3SG

‘S/he again didn’t give it to her/him.’

In Digor, the placement of a clitic cluster is somewhat less restricted, although the acceptability of a sentence decreases when the cluster is shifted rightwards. Clitics in a cluster are subject to certain restrictions on case-person combinations. The cluster is normally rigidly ordered, but some speakers allow reordering to improve certain combinations of case and person. Modern Ossetic does not allow clitic doubling, unless the doubled noun phrase is on the right edge of the clause. Clitics may be hosted by non-finite clauses as well, (14).

- (14) [*enqelme=dəm keš-ge-je] ba-fənej dən* (I)
 waiting=ALL.2SG look-CVB-ABL PRV-sleep be.PRS.1SG
 'I fell asleep while waiting for you.'

4.6. Questions

The word order in yes-no and alternative questions does not differ from that in declarative clauses. In wh-questions, all wh-phrases are obligatorily preverbal with only second position clitics, negative indefinites, negation markers, and certain adverbials being able to intervene between the wh-complex and the verb, see details in Erschler (2012a).

4.7. Negation

Sentential negation is marked by mood-dependent proclitics to the verb. Negative indefinites are all placed in the immediately preverbal position. They cannot co-occur with sentential negation markers, see details in Erschler and Volk (2011). Unlike in other Ossetic varieties, the South Ossetian Kudar dialect, obviously under recent influences from Georgian, optionally allows for negative indefinites to co-occur with sentential negation:

- (15) *nišə (næ) fed-ton¹⁶* (Kudar Iron)
 nothing NEG see.PST-PST.1SG
 'I saw nothing.'

4.8. Binding

Reflexives and reciprocals can in principle be bound by any clause-mate NP, although the subject is the preferred binder. Pronominal enclitics are free in the minimal clause. Possessive proclitics are free from binding restrictions. Full pronouns, including full possessive pronouns, are by default free in the sentence, (16a). However, if the full pronoun is focused or topicalized, it can have an antecedent within the same sentence, (16b).

- (16) a. *soslan-me_i wote kes-uj cuma je_{j/?i}* (D)
 Soslan-ALL so look-PRS.3SG COMP (s)he
me=zerde-me cəw-uj
 poss.1SG=heart-ALL go-PRS.3SG
 'Soslan_i thinks that I like him_{j/?i}'

¹⁶ The transcription represents the Kudar pronunciation.

- b. *medine* *erme* *woj_{i/ɔj}* *ke* *warz-uj* *wote*
 Madina only s/he.OBL COMP love-PRS.3SG so
soslan *buđi* *ken-uj* (D)
 Soslan thought do-PRS.3SG
 ‘Soslan_i thinks that Madina only loves him_{i/ɔj}’

5. Clause linkage

5.1. Coordination

The conjunctions *eme/ema* ‘and’ and *fele/fal* ‘but’ are clause-initial, second position clitics then attach to conjunctions, like the Iron *jem* ALL.3SG in (17b):

- (17) a. *nər* *vender* *ran* *ba-lxəd-ta* *big^{wəz}* *χəzar*
 now other place PRV-buy.PST-PST.3SG Bigwyz house
eme *adon* *wəj* *ken-ə* (I)
 and these_ones.NOM sale do-PRS.3SG
 ‘Now, Bigwyz has bought a house in a different place and is selling
 these ones.’
 (K’æbysty 1977)
- b. *mən=dər* *jv=χi-mə* *ragej* *ždeχta*
 I.OBL=too POSS.3SG=REFL-ALL since.long attract.PST.3SG
fantastike *fele=jem* *mə=nəʃ* *ne*
 sci-fi but=ALL.3SG POSS.1SG=certainty NEG
χašton carry.PST.1SG (I)
 ‘The (genre) of science fiction has long attracted me too, but I did
 not dare to approach it.’
 (K’æbysty 1977)

Noun phrases (but not clauses) can also be coordinated by encliticizing *dər* to each one of them, (18a). The usual conjunction ‘and’ can also be used in the presence of *dər*, (18b).

- (18) a. *χ^warz-vj=dər* *cud-vj=dər* *či* *fe-wwid-ton* (D)
 good-ABL=and bad-ABL=and what PRV-see.PST-PST.1SG
 ‘What I have seen of good and of bad’
 (Aghuzarti 2008)
- b. *č’ife=dər* *ema* *wazal=dər* (D)
 dampness=and and cold=and
 ‘dampness and cold’
 (Aghuzarti 2008)

There are two conjunctions for ‘or’: *kene/kene* is used when the two conjuncts are not mutually exclusive, and *evi* in all other situations. They are placed clause-initially (or constituent-initially, if constituents smaller than clauses are coordinated):

- (19) a. *ez kiwunuge fins-un ne=kav-un kene ender*
 I.NOM book write-INF NEG=intend-INF or other
ješti woj xuzen
 something it.OBL similar
 ‘I am not going to write a book or anything like that.’
 (Aghuzarti 2008)
- b. *də=məm evəg zur-əš evi=mə mə=quš-tə*
 you=ALL.2SG real talk-PRS.2SG or=ACC.1SG POSS.1SG=ear-PL
šaj-ənc?
 deceive-PRS.3PL
 ‘Are you really speaking to me or do my ears deceive me?’
 (K’æbysty 1977)

5.2. Subordination

All subordinate clauses allow or even require a proleptic noun phrase in the main clause. The conditions under which proleptics are obligatory are not yet well understood. In (20), proleptics are marked with boldface.

- (20) a. *[səmej žonənad ba-nk’ar-aj] wəj təx̥xəj*
 in_order_that science PRV-feel-SUB.FUT.2SG it because
ra-jdaj iwwəl enson-dər-ej
 PRV-begin.IMP.2SG it.SUP easy-COMP-ABL
 ‘In order to internalize the knowledge, you should start from the easiest.’
 (K’æbysty 1977)
- b. *ez nisə ažəmdžən dən [zur-ən kej*
 I nothing guilty be.PRS.1SG talk-INF COMP
žon-ən eme matematikə aſte xorž kej š-aχʷər
 know-PRS.1SG and mathematics so good COMP PRV-learn
kod-ton] wəm
 do.PST-PST.1SG it.INES
 ‘I am not at all guilty that I can speak and have learned mathematics so well.’
 (K’æbysty 1977)

Complex clauses do not show any consequence of tense effects, whereby the tense of the verb in the main clause would impose restrictions on the tense marking in the dependent clause. In (21), the past tense of the main verb does not block the future marking in the dependent clause.

- (21) *šošlan-mv* *aftə* *keš-ə* *kašt-iš* *vme* *medine*
 Soslan-ALL so look-PRS.3SG look.PST-PST.3SG and Madina
kej *erba-sewu-zəniš* (I)
 COMP PRV-go-FUT.3SG
 ‘Soslan thinks/thought that Madina will/would come.’

5.3. Complementation

Complementizers are either strictly preverbal with same qualifications as for *wh*-phrases as described in §4.6, (22a), or float between the left edge of the clause and the verb, Erschler (2012a), (22b). The complement clause usually follows the matrix verb, but may precede it as well. A finite dependent clause is typically accompanied by a proleptic in the main one.

- (22) a. *ragej* *χat-ə* *[artur vme gertrud eŋə*
 since.long notice-PRS.3SG Arthur and Gertrude without
kerezi *kej* *nə=fe-raž-ənc]* *wəj* (I)
 each_other COMP NEG=PRV-withstand-PRS.3PL it
 ‘Since long, she’s noticed that Arthur and Gertrude cannot live without
 each other.’
 (K’æbysty 1977)
- b. *felə* *də=žardə* *ked vender-en wa*
 but POSS.2SG=heart if other-DAT be.SUB.FUT.3SG
leverd (I)
 give.PRT.PST
 ‘But if your heart would be given to someone other’
 (K’æbysty 1977)

Some verbs allow a null complementizer, as illustrated in (23); *wəj* it.NOM/OBL serves here as a proleptic:

- (23) *l_{cp}-a-izer=ma* *šeri-jə* *fen-zən]* *l_{MATRIX}wəj enqəl*
 this-evening=more Seri-OBL see-FUT.3SG it.NOM/OBL hope
nal *wədi]* (I)
 no_more be.PST.3SG
 ‘She already did not think that she would see Seri tonight.’
 (Bic’oty 2003)

The coordinating conjunction *emə / ema* ‘and’ has grammaticalized into a subordinator:

- (24) a. *bazel-ə qʷədə-tə aftə nə-ššuj-tə štə*
Bazel-OBL thought-PL so PRV-gut-PL be.PRS.3PL
emə=zə rajdajen-keron bərəg nal wəd
and=ABL.3 beginning-end sign no_more be.PST.3SG
‘Bazel’s thoughts got so entangled that it was not possible to find their beginning or end.’
(K’æbysty 1977)
- b. *ənqəl dən emə mə=qədəz-ə*
thought be.PRS.1SG and POSS.1SG=grief-OBL
nə=ba-səw-zənə
NEG=PRV-go-FUT.2SG
‘I think/hope that you won’t be angry with me?’
(K’æbysty 1977)

Non-finite verb forms may head subordinate clauses as well, for instance, purpose clauses are headed by infinitives:

- (25) *doxtur-tə=min bərə ra-vardt-once [kust-mə*
doctor-PL=DAT.1SG right PRV-give.PST-PST.3PL work-ALL
cəw-un]-mə go-INF-ALL
‘Doctors permitted me to go to work.’
(Aghuzarti 2008)

5.4. Relative clauses and multiple relativization

A noun phrase can be relativized from any syntactic position. Relative pronouns, which are homophonous with wh-words, are immediately preverbal, with the same qualifications as for wh-phrases. Relative clauses are formed using the correlative strategy: the head noun – the correlate – follows the relative clause and must carry the distal demonstrative. In (26), the relative pronouns and the correlates are highlighted in boldface:

- (26) a. *parč-ə sə čəžg-imə nəχaš kod-ta*
park-OBL what girl-COM talk do.PST-PST.3SG
wəj-imə jə=k'am ſ-išta
s/he-COM POSS.3SG=picture PRV-take-PST.3SG
‘He had his picture taken with the girl that he spoke with in the park.’

- b. *čožg-imə kem nəχaš kod-ta wəsə parč-ə*
 girl-COM where talk do.PST-PST.3SG that parc-OBL
je=k'am š-iš-ta
 POSS.3SG=picture PRV-take.PST-PST.3SG
 ‘He had his picture taken in the park where he talked with the girl.’
- c. *wədon še=institut-mə aχʷər-mə ra-jš-toj*
 they POSS.3PL=institute-ALL learning-ALL PRV-take-PST.3PL
[siri-ja či ser-ə] wəsə kəšgon leppu-t-ə
 Syria-OBL who live-PRS.3SG that Circassian boy-PL-OBL
vme čoždž-ət-ə k'ord-ə (I)
 and girl-PL-OBL group-OBL
 ‘They accepted to study in their institute a group of Circassian
 boys and girls who live in Syria.’¹⁷

Given that quantifiers are not very well compatible with demonstratives, to relativize a universally or existentially quantified noun phrase, one needs to use different strategies, namely, free relatives or non-finite dependent clauses:

- (27) a. *[čideritter xəreg dar-ə] [alčider=əj səv-gə*
 whoever donkey own-PRS.3SG everyone=ACC.3SG beat-CVB
kən-ə] (I)
 do-PRS.3SG
 ‘Anyone who owns a donkey, beats it.’ (Lit. ‘Whoever owns a donkey,
 everyone beats it.’)
- b. *[dərkəvš-mə səw-ag išci] gošt-mə akumul'ator*
 Darghavs-ALL go-NMZ someone Gosht-ALL accumulator
a-laš-zən (I)
 PRV-carry-FUT.3SG
 ‘Someone who is going to Darghavs will take the accumulator to
 Gosht.’

It is also possible here to place the quantified noun phrase sentence-initially and to use it as a binder for the regular demonstrative:

- (28) *alə žəχxkušag=der [xəreg či dar-ə] wəj=jə*
 every peasant=EMP donkey who own-PRS.3SG he=ACC.3SG
səv-gə=der kən-ə (I)
 beat-CVB=EMP do-PRS.3SG
 ‘Every peasant who owns a donkey beats it.’

Ossetic allows multiple relativization, relative pronouns are then all placed preverbally. Such clauses only allow for generic reading:

¹⁷ (<http://www.iriston.com/nogbon/news.php?newsid=543>, accessed 08 July 2017).

- (29) a. *kem=ju kejime nəχaš kod-ta wəm*
 where=HAB who.COM talk do.PST-PST.3SG there
alkejime je=k'am išta (I)
 everyone.COM POSS.3SG=picture take-PST.3SG
 ‘Wherever s/he spoke with anyone, s/he had his picture taken with
 them.’
- b. **kem sə čəžg-imə nəχaš kod-ta wəsə parč-ə*
 where what girl-COM talk do.PST-PST.3SG that park-OBL
wəsə čəžg-imə je=k'am š-išta (I)
 that girl-COM POSS.3SG=picture PRV-take-PST.3SG
 Intended: ‘When he spoke to a girl in a park, he had his
 picture taken with her.’¹⁸

6. Language Contact Issues

The typological profile of Ossetic, as it has been described in this chapter, is fairly different from those of other modern Iranian languages. It is natural to conjecture that at least some of these properties arose due to language contact. It should be kept in mind, however, that Ossetians and their ancestors have long been isolated from the rest of Iranian speakers. The latter occupy a more or less contiguous territory and have been significantly influenced by Persian language and culture. Some of the differences could have emerged, therefore, just as independent developments.

One instance of Caucasian influence is fairly obvious: this is the most natural way to explain the emergence of ejectives in Ossetic. While virtually ubiquitous in the Caucasus, where their presence is essentially the only shared areal feature, Tuite (1999), ejective consonants are fairly uncommon in the rest of Eurasia and are completely absent from other Iranian languages.

The fact that Ossetic extensively borrowed vocabulary from the neighboring languages is also well-documented, see Abaev (1958–1995), Bielmeier (1977), and Thordarson (1999), to name a few.

Contact influences in the realm of morphosyntax are admittedly harder to argue for. Given that data on earlier stages of Ossetic are virtually absent, the only way to prove that a given grammatical feature is a result of borrowing or influence from a specific source is to show that (a) the feature in question is present in Ossetic and in the source language(s); (b) that the given feature is rare cross-linguistically, or, at least in the relevant area – i. e. northern Eurasia including the Caucasus.

One grammatical feature often claimed to be an outcome of usually unspeci-

¹⁸ It should be emphasized that respective single relatives are fully grammatical, as illustrated in (26).

fied Caucasian influences is the relatively large Ossetic case system, (Vogt 1945; Abaev 1949: 99). Sometimes, specifically South Caucasian influences are postulated, Johanson (2008: 500). It is certainly likely that the emergence of the case system, abnormally large for a Modern Iranian language, was an outcome of some influence(s). However, neither the size, nor the case inventory are particularly unusual for northern Eurasia, Erschler (2009: 423), Thordarson (2009: 170). The fact of Caucasian influences on this system is not implausible, but hardly demonstrable with any degree of certainty. Moreover, one salient feature of many Caucasian languages – namely, the presence of ergative marking, is entirely absent from Ossetic.

It seems more promising to conjecture that possessive proclitics, the structure of wh-questions, and the morphosyntax of negative indefinites were influenced by neighboring autochthonous languages. Possessive proclitics, unique for Iranian languages and in general very rare in Eurasia, have been argued to be a result of West Caucasian influences, Erschler (2009). Obligatory preverbal placement of wh-phrases is likely to be a South Caucasian influence, Erschler (2012). Finally, the morphology of negative indefinites and their incompatibility with sentential negation is likely to be due to South Caucasian influences as well, Erschler (2010).

Interestingly, no instances of potential grammatical influences from Northeast Caucasian languages have been discovered so far. However, there is one putative influence in the opposite direction: emergence of phonemic /ʃ/ in Ingush was attributed to Ossetic influence already by Uslar (1888: 6). Uslar himself, to support this conjecture, only observes that Chechen lacks /ʃ/ and states that “their (i. e. Ingush) vernacular was formed under a strong influence of Ossetic language”. The supporting argument is that the phonemic /ʃ/ is absent from almost all the Northeast Caucasian languages, and from the South Caucasian languages. This is the only known instance of a plausible Ossetic influence on a neighboring language.

Another potential instance of contact induced development is a two-term deictic system present in Ossetic, Svan, Mingrelian, and Karachay-Balkar – a fact noticed by Thordarson (2009). However, another Eastern Iranian language, Yazghulami, has developed a two-term system without obvious external influences.

Preverbs with spatial semantics are present in virtually all autochthonous languages of the Caucasus, but it is only in the modern SC languages and in Ossetic that they (almost) always have a perfectivizing function. It is so in Georgian (Boeder 2005: 33), as well as for most preverbs in Mingrelian, Harris (1991), and in Svan, Tuite (1997: 30). Such systems of perfectivizing preverbs strikingly resemble the Slavic one. Furthermore, Abaev (1965: 60–62) argued that it is the ancestors of Ossetians came in contact with the Slavs and brought the innovation to the Caucasus. However, Thordarson (1982: 252–253) convincingly argues against this conjecture. See also the discussion in Tomelleri (2009).

Finally, Mingrelian has an imperfectivizing affix placed in a verb between the prefixes and the root (Harris 1991), a phenomenon similar to the Ossetic conative

suffix *-sej-/ -cej-*, but this parallel might well be due to chance: for one thing, direct contacts between Ossetians and Mingrelians are not documented nor are they very likely for geographical reasons.

For the last 150 years, Ossetic has been under strong Russian influence. Code-switching is very frequent and virtually any Russian word can be occasionally used in Ossetic speech.

7. Short glossed text

A sample glossed Digor oral text. Recorded 20 July, 2008 in Sheker, Iraf district, North Ossetia from Mimonat Gogaeva. Transcribed in the summer of 2009 with help of Aza Sasieva.

- (30) *men-en encon-der wodzej digoron-aw džor-un*
 I.OBL-DAT easy-COMP be.FUT.3SG Digor-EQU speak-INF
 'It will be easier for me to speak in Digor'

- (31) *ez lešken-i ra-jgurd-ten mijn farast*
 I Lesken-OBL PRV-born.PST-PST.1SG 1000 9
sede duwin ertikkag anž-i
 100 20 third year-OBL
 I was born in Lesken, in the year 1923. [A decimal numeral.]

- (32) *ema lešken-i adtej hewežen eždewu-tte*
 and Lesken-OBL be.PST.3SG such custom-PL
 And in Lesken there were the following customs. [Note the singular verb with the plural of 'custom'.]

- (33) *ficcagi-der xester-en kade levard-tonce xester ku*
 first-COMP senior-DAT respect give.PST-PST.3PL senior when
fe-štide wed ežwame kester-te
 PRV-stand_up.PST-SUB.PST.3SG then needs junior-PL
jewugur-ej=der iš-istad-ajonce
 together-ABL=EMP PRV-stand_up.PST-SUB.PST.3PL
 First of all, they respected seniors, if an elder would stand up, then all the juniors would have to stand up as well.

- (34) *xester-en ežwame eždaw levard-tajonce*
 elder-DAT needs etiquette give.PST-SUB.PST.3PL
 (The juniors) would have to behave courteously (according to the etiquette) towards the seniors.

- (35) *kester-en dzurd-i bare n=adtej calimme xester*
 junior-DAT word-OBL right NEG=be.PST.3SG while elder
dzurd-i bare ravard-tajde wed-me
 word-OBL right give.PST-SUB.PST.3SG then-ALL
 To a junior there would be no right to speak, until the elder gives him right to speak.
- (36) *xester ka=j woj kud fende wide*
 senior who=be.PRS.3SG s/he.OBL how wish be.SUB.PST.3PL
wote ƿewame kod-tajonce kester-te
 so needs do.PST-SUB.PST.3PL junior-PL
 As the elder's wish would be, so should act the juniors.
- (37) *woj fendon enxest kod-tajonce*
 s/he.OBL will complete do.PST-SUB.PST.3PL
 They would have to fulfill his/her wishes.
- (38) *berəχ̥arž-dzijned-te adtej ragi nur=səvel*
 many good-ABSTR-PL be.PST.3SG earlier now=3SG.SUP
sə=k'ax-t-ej fe-χatt-encə
 POSS.3PL=leg-PL-ABL PRV-tread.PST-PST.3PL
 Originally, there was much good, now they tread upon it with their feet.
- (39) *ma woni ne-bal imiš-dzinan fal=sə imis-un*
 and they.OBL NEG-more remember-FUT.1PL but=ACC.3PL remember-INF
ƿew-uj newag-ej=sə er-lew-un ƿew-uj
 need-PRS.3SG new-ABL=ACC.3PL PRV-stand-INF need-PRS.3SG
 We do not remember them now, but we must remember them and reinstate them. [In principle, it should be the causative construction, *er-lew-un ken-un ƿew-uj* PRV-stand-INF do-INF need-PRS.3SG, but *ken-un* may sometimes be dropped in the casual speech.]
- (40) *erdaŋ kemi ne wa womi rasuwd-dzijnadə*
 etiquette where NEG be.SUB.FUT.3SG there beautiful-ABSTR
ne-ked wodženej ne-ked
 NEG-when be.FUT.3SG NEG-when
 Where there will be no etiquette, never will there be any beauty, never.'

Abbreviations

ABL	ablative	IMP	imperative
ABSTR	abstract noun derivation	INES	inessive
ACC	accusative	INF	infinitive
ALL	allative	NEG	negative
AOR	aorist	NMZ	nominalizer
ASS	associative plural	NOM	nominative
COM	comitative	NUM	numeral declension
COMP	complementizer	OBL	oblique
CON	conative	PL	plural
COP	copula	POSS	possessive
CTR	contrastive topic	PRS	present
CVB	converb/gerund	PRT	participle
DAT	dative	PRV	preverb
EMP	emphatic clitic	PST	past
EQU	equative	SG	singular
FUT	future	SUB	subjunctive
HAB	habitual	SUP	supercessive
IDF	indefinite		

References

References to the sources of examples

- Aghuzarti, Agube. 2008. *Zin cardi k'æpxæn* [Stairs of Hard Life]. Dzæwugighæu: CI NEI RPX.
- Bedzhyzaty, Chermen. 1995. *Uacmystæ* [Stories]. Dzæuydžyqæu: Ir.
- Bic'oty, Griš. 2003. *Kærdædžy st'aly* [The Evening Star]. Dzæuydžyqæu: Ir.
- Dzasoxti, Muzafer. 2007. *Dællag Ir* [Lower Ossetia]. Dzæuydžyqæu: Ir.
- Gurdzibety, Baris. 2006. *Uarzt æma fudænxæ* [Love and hate]. Dzæugighæu: Olimp.
- K'æbysty, Zaur. 1977. *Fæstag majmuli* [Last monkey]. Cxinval: Iryston.
- Qajtyqty, Azæmæt. 1998. *Xod: Taurægh æmæ card* [Khod: Legend and life]. Dzæuydžyqæu: Ir.
- Sabajti, Sulejman. 2010. *Radzurdtæ* [Stories]. Dzæugighæw: Gassiti Viktori nombæl rauaghdon-poligrafion kustuat.

Other references

- Abaev, Vasilij. 1949. *Osetinskij jazyk i fol'klor* [Ossetic language and folklore]. Moscow & Leningrad: Izdatel'stvo AN SSSR.
- Abaev, Vasilij. 1958/1973/1979/1989/1995. *Istoriko-etimologičeskij slovar' osetinskogo jazyka* [Historical etymological dictionary of Ossetic]. 5 vols. Moskva: Izdatel'stvo Akademii Nauk SSSR.
- Abaev, Vasilij. 1964. *A grammatical sketch of Ossetic*. Den Haag: Mouton.

- Abaev, Vasilij. 1965. *Skifo-evropejskie izoglossy* [Scythian-European isoglosses]. Moscow: Nauka.
- Abaev, Vasilij & H. W. Bailey. 1985. Alans *Encyclopædia Iranica* 1.8, 801–803. Winona Lake, IN: Eisenbrauns.
- Alemany, Augusti. 2000. *Sources on the Alans: A Critical Compilation*. Leiden & Boston: Brill.
- Axvlediani, Georgij (ed.). 1963/1969. *Grammatika osetinskogo jazyka* [A grammar of Ossetic]. 2 vols. Ordžonikidze: Severo-Osetinskij Naučno-Issledovatel'skij Institut.
- Bailey, H. W. 1950. L'accento in osseto digoron. *Ricerche Linguistiche* I. 58–69.
- Bagaev, Nikolaj. 1965/1982. *Sovremennyj osetinskij jazyk* [Modern Ossetic], part 1, 2. Ordžonikidze: Ir.
- Bekoev, Dmitri. 1985. *Ironskij dialekt osetinskogo jazyka* [The Iron dialect of Ossetic]. Cxinval: Iryston.
- Belyaev, Oleg & Arseniy Vydrin. 2011. Participle-converbs in Iron Ossetic: Syntactic and Semantic Properties. In Geoffrey Haig, Simin Karimi, Agnes Korn & Pollet Samvelian (eds.), *Topics in Iranian Linguistics* (Beiträge zur Iranistik Bd. 34), 117–134. Wiesbaden: Reichert.
- Bielmeier, Roland. 1977. *Historische Untersuchung zum Erb- und Lehnwortschatzanteil im ossetischen Grundwortschatz*. Frankfurt am Main: Lang.
- Boeder, Winfried. 2005. The South Caucasian languages. *Lingua* 115. 5–89.
- Christensen, Arthur Emanuel. 1921. *Textes ossètes* (Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser 6/1). Copenhagen: Publisher.
- Dzaxova, Veronika. 2009. *Fonetičeskie xarakteristiki fonologičeskoy sistemy sovremenennogo osetinskogo (ironskogo) literaturnogo jazyka* [Phonetic Characterization of the Phonological System of the Standard Modern Iron Ossetic]. Vladikavkaz: Izdatel'stvo SOGPI.
- Dzaxova, Veronika. 2010. Ob osetinskom udarenii [On the Ossetic Stress]. *Vestnik RGGU* 9/10(52). 9–26.
- Dzodzikova, Zaida. 2009. *Analitizm v slovoobrazovanií osetinskogo jazyka* [Analyticity in the Ossetic Word Formation]. Vladikavkaz: SOIGSI.
- Erschler, David. 2009. Possession marking in Ossetic: Arguing for Caucasian influences. *Linguistic Typology* 13(3). 417–450.
- Erschler, David. 2010. Otricatel'nye mestoimenija v osetinskom jazyke: areal'nye i tipologicheskie aspekty [Negative Pronouns in Ossetic: Areal and Typological Issues]. *Voprosy Jazykoznanija* 2010(2). 84–105.
- Erschler, David. 2012a. From preverbal focus to preverbal “left periphery”: On diachronic and typological aspects of the Ossetic clause structure. *Lingua* 122. 673–699.
- Erschler, David. 2012b. Suspended Affixation and the Structure of Syntax-Morphology Interface. *Studia Linguistica Hungarica* 59. 153–175.
- Erschler, David. 2014. A marginal modal form in Diger Ossetic. *Acta Linguistica Petropolitana* X(3). 711–719.
- Erschler, David. 2016. Ossetic. In Peter O. Mueller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation: An International Handbook of the languages of Europe*, 3157–3179. Berlin & Boston: De Gruyter Mouton.
- Erschler, David & Vitaly Volk. 2011. On negation, negative concord, and negative imperatives in Diger Ossetic. In Geoffrey Haig, Simin Karimi, Agnes Korn & Pollet Samvelian (eds.), *Topics in Iranian Linguistics* (Beiträge zur Iranistik Bd. 34), 135–152. Wiesbaden: Reichert.

- Gabaraev, Nikolaj. 1977. *Morfologičeskaja struktura slova i slovvobrazovanie v osetinskem jazyke* [Morphological word structure and word formation in Ossetic]. Tbilisi: Mecniereba.
- Harris, Alice C. 1991. Mingrelian. In Alice C. Harris (ed.), *The Indigenous Languages of the Caucasus*, vol. 1, 313–394. Delmar, NY: Caravan Books.
- Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.
- Isaev, Magomet I. 1966. *Digorskij dialekt osetinskogo jazyka* [The Digor dialect of Ossetic]. Moskva: Nauka.
- Johanson, Lars. 2008. Case and contact linguistics. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 494–501. Oxford: Oxford University Press.
- Kambolov, Tamerlan. 2007. *Jazykovaja situacija i jazykovaja politika v Severnoj Osetii: Istorija, sovremennost', perspektivy* [Language situation and language policy in North Ossetia: History, modernity, perspectives]. Vladikavkaz: Ir.
- Kudzoeva, Andžela. 2003. *Porjadok slov i aktual'noe členenie v osetinskom jazyke* [Word Order and Communicative Structure in Ossetic]. Vladikavkaz: SOGU.
- Miller, Wsewolod. 1881. *Osetinskie etudy. Chast' I. Osetinskye teksty* [Ossetic Sketches. Part 1. Ossetic Texts]. Moscow: Imperatorskij Moskovskij Universitet.
- Miller, Wsewolod. 1903. *Die Sprache der Osseten*. Strassburg: Trübner.
- Miller, Wsewolod & Reinhold von Stackelberg. 1891. *Fünf ossetische Erzählungen in digorischem Dialect*. Mit deutscher Übersetzung, Glossar und Anhang von R. von Stackelberg. St. Petersburg: Académie Impériale des Sciences.
- Munkácsi, Bernhard. 1932. Blüten der ossetischen Volksdichtung. *Keleti Szemle* 21. 68–75.
- Németh, Gyula. 1959. Eine Wörterliste der Jassen, der ungarländischen Alanen. *Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin. Klasse für Sprachen, Literatur und Kunst* 1958(4). Berlin: Akademie-Verlag.
- Rosen, Georg. 1846. *Ossetische Sprachlehre*. Berlin: Lemgo und Detmold.
- Sims-Williams, Nicholas. 1996. Eastern Iranian Languages. *Encyclopædia Iranica* 7.6, 649–652. New York: Center for Iranian Studies, Columbia University. <http://www.iranicaonline.org/articles/eastern-iranian-languages> (accessed 23 September 2017).
- Sjögren, Andreas Johan. 1844. *Ipon ሂይզያኬም das ist Ossetische Sprachlehre. Nebst kurzem ossetisch-deutsche und deutsch-ossetische Wörterbuche*. St. Petersburg: Kaiserliche Akademie der Wissenschaften.
- Sokolova, Vera. 1953. *Očerki po fonetike iranskix jazykov* [Sketches of Iranian Phonetics], part 2. Moscow-Leningrad: Izdatel'stvo AN SSSR.
- Stackelberg, Reinhold von. 1886. *Beiträge zur Syntax des Ossetischen*. Strassburg: Trübner.
- Takazov, Fedar. 2003. *Digor-urussag dzurduat* [Digor-Russian dictionary]. Vladikavkaz: Alanija.
- Takazov, Fedar. 2009. *Grammatičeskij očerk osetinskogo (digorskogo) jazyka* [A grammar sketch of Digor Ossetic]. Vladikavkaz: Izdatel'stvo SOGU.
- Tedeev, Geras. 1989. *Semantickaja struktura osetinskogo glagola* [The Semantic Structure of the Ossetic Verb]. Tbilisi: Mecniereba.
- Tekhov, Fedor. 1970. *Výraženie modal'nosti v osetinskem jazyke* [Expression of modality in Ossetic]. Tbilisi: Mecniereba.
- Testen, David. 1997. Ossetic Phonology. In Alan S. Kaye (ed.), *Phonologies of Asia and Africa: Including the Caucasus*, vol. 1, 707–732. Winona Lake, Indiana: Eisenbrauns.
- Thordarson, Fridrik. 1982. Preverbs in Ossetic. In *Monumentum Georg Morgenstierne*, vol. 2, (Acta Iranica 22), 251–261. Leiden: Brill.

- Thordarson, Fridrik. 1999. Linguistic contacts between the Ossetes and Kartvelians: A few remarks. In Helma van den Berg (ed.), *Studies in Caucasus linguistics: Selected papers of the Eighth Caucasian Colloquium, Research School of Asian, African and Amerindian Studies (CNWS)*, 279–285. Leiden: Universiteit Leiden.
- Thordarsson, Fridrik. 2009. *Ossetic grammatical studies* (Sitzungsberichte der philosophisch-historischen Klasse, Bd. 788, Veröffentlichungen zur Iranistik, Nr. 48). Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Tomelleri, Vittorio. 2009. The Category of Aspect in Georgian, Ossetic and Russian. Some areal and typological observations. *Faits de langues* 2009(1). 245–272.
- Tuite, Kevin. 1997. *Svan* (Languages of the World/Materials 139). München: LINCOM Europa.
- Tuite, Kevin. 1999. The myth of the Caucasian Sprachbund: The case of ergativity. *Lingua* 108. 1–29.
- Uslar, Petr. 1888. *Čečenskij jazyk* [Chechen]. Tiflis: Izdanie Upravlenija Kavkazskago Učebnago Okruga.
- Vogt, Hans. 1945. Substrat et convergence dans l'évolution linguistique. Remarques sur l'évolution et la structure de l'arménien, du géorgien, de l'ossete et du turc. *Studia Septentrionalia* 2. 213–228.
- Volkova, Natalya. 1974. *Etničeskij sostav naselenija Severnogo Kavkaza v XVIII – načale XX veka* [The ethnic composition of the population of the North Caucasus in the 18th – early 19th centuries]. Moscow: Nauka.

6.4. Romeyka

Laurentia Schreiber

1. Introduction

Romeyka is an endangered Indo-European language, belonging to the Hellenic branch. It is still spoken *in situ* by Muslim communities on the southern Black Sea coast in Turkey, and by diaspora speakers in Turkey's large cities and elsewhere (see §1.3). A complex set of geographical and historical circumstances have led to its current sociolinguistic situation, which is outlined in the initial sections of this chapter. Many structural aspects of Romeyka are still only poorly understood, though ongoing research is rapidly filling many of the gaps in our knowledge (Sitaridou 2013, 2014a, 2014b, and references therein). This chapter summarizes the current state of knowledge, and refers to the available literature for more detailed coverage.¹ The sociolinguistic information in this chapter is based on Schreiber (2016), Schreiber and Sitaridou (2017).

1.1. Naming conventions

Romeyka is the term that has become widespread in recent research,² though “Muslim Pontic”, a term introduced by Mackridge (1987), is used as a technical term, too (namely by Özkan 2013; partially Brendemoen 2006; Bortone 2009; see Schreiber and Sitaridou 2017 on glossonomy). The term Romeyka³ is in line with what many speakers call their language, though others use the term “Rumca”, especially when speaking in Turkish. As argued by Schreiber and Sitaridou (2017), the use of either the Greek term *Romeyka*, or the Turkish term *Rumca*, corresponds to the ethnolinguistic vitality of the particular speech community. Regional variants are referred to with name of the location in which the variety is spoken (e. g. Romeyka of Of, Romeyka of Sürmene). When referring to other varieties of Pontic

¹ I wish to express my sincere gratitude to Dr Hakan Özkan for sharing his data from Romeyka of Sürmene. I thank Prof Geoffrey Haig for his rich feedback and insightful discussion of this chapter. Moreover, I am deeply indebted to the Istanbulite Rumca speaker who generously and patiently provided me an insight into her language. I dedicate this chapter to her. Furthermore, I thank Dr Kilu von Prince for her support in the early stage of this chapter. I acknowledge that all errors are my own.

² Cf. in particular work undertaken in the British Academy Project ‘Description and Documentation of the Romeyka varieties in Pontus; Continuity, Contact and Change’, PI: Dr Ioanna Sitaridou, and consequent publications; www.romeyska.org.

³ For variation in spelling see <http://www.romeyska.org/the-romeyska-project/rediscovering-romeyska> (accessed 07 August 2017).

Greek, the following terms are used: the term “Pontic Greek” without any supplement denotes the language of Christian Pontic Greek speakers in Greece after 1923; “Christian Pontic Greek” is used for the Christian Pontic Greek speakers in Pontus until 1923.

1.2. Language classification

Romeyka belongs to the Hellenic branch of the Indo-European language family. It is part of the Asia Minor Greek group where it is considered to form a sub-group with Cappadocian as opposed to other Asia Minor Greek varieties such as Siliot (cf. Sitaridou 2014b). Several aspects of Romeyka’s origins remain controversial: it is unclear how far Romeyka was involved in the development of an Asia Minor *koine* (Sitaridou 2014b: 29), and whether Medieval Pontic had diverged into Christian and Muslim varieties already after Islamization in the 16th/17th century, and thus before expulsion of the Christian speakers from Turkey in 1923, as argued by Sitaridou (2014b: 31).

1.3. Sociolinguistic vitality

Assessments of the sociolinguistic vitality of Romeyka vary: “definitely endangered” (Moseley 2010), “seriously/severely endangered” (Moseley 2007), “threatened” (Catalogue of Endangered Languages 2017) and “vigorous” (Lewis, Simons, and Fennig 2016). These assessments are questionable because they (a) do not differentiate between Romeyka and Pontic Greek (hence speaker numbers are inaccurate), and (b) are based on insufficient data (Schreiber and Sitaridou 2017).

The actual number of Romeyka speakers is not known. The last official number reports 4,535 speakers and stems from the 1965 Turkish general census (*Genel Nüfus Sayımı*), which was the last of its kind accounting for different mother tongues (reported in Mackridge 1987; Andrews 1989). In general, the estimation of speaker numbers is difficult due to (a) practical reasons such as internal and external migration and residence patterns, (b) a biased choice of Turkish as mother tongue, and (c) mixed Turkish-Romeyka identity at least in the urban communities. Estimates based on the populations of the villages in Trabzon province where Romeyka is spoken run to approx. 300,000 speakers (Bilici 2011), but this kind of estimate is obviously exceedingly coarse-grained (see Schreiber and Sitaridou 2017).

Romeyka is spoken in approximately fifty villages in Trabzon province around the cities of Sürmene, Of, Çaykara and Tonya (Mackridge 1987; Andrews 1989; Özkan 2013). Furthermore, since labour migration took place as of the 1950s, there exist speech communities in urban centres such as Ankara and Istanbul whose language vitality differs significantly from rural communities (Schreiber and Sitaridou 2017). Since 1960, diaspora communities exist in Germany and other central

European countries (Özkan 2013). A considerable threat to Pontic Greek as spoken by Christians in the Black Sea region was the relocation of Christian Pontic Greek speakers in Greece in the context of the Treaty of Lausanne in 1923. However, already before the 20th century, Pontic Greeks had migrated to southern Russia and Caucasia (Moseley 2007) and later to the US, the Ukraine, and Greece (Brendemoen 2002).

Romeyka speakers may harbour multiple identity constructs, such as Turkish national and citizenship identity, Muslim religious identity and, dependent on the speech community, a sense of cultural Romeyka identity (Sitaridou 2013). The latter is less pronounced in urban communities, which tend towards assimilation into a Turkish mainstream (Schreiber 2016). In general, Romeyka speakers deny any Rum or Greek identity (Sitaridou 2013) and often exhibit conservative Muslim and nationalistic identification links (Özkan 2013).

All Romeyka speakers are bilingual in Romeyka and Turkish. Depending on the speech community, the youngest generation acquires Romeyka still as early L2 or, respectively, as heritage language only (Schreiber and Sitaridou 2017).

The development of a written form of Romeyka is not encouraged by Turkish language education policies, and it generally lacks the range of registers for more official domains. As a result (and a cause), its use is restricted to informal domains, and functions now as home and heritage language, with an ongoing decline (Schreiber 2016). So far, no widely-accepted orthography has been developed for Romeyka, and we lack a comprehensive descriptive grammar or dictionary, though documentation is underway (see Fn. 1).

1.4. Dialectal variation

Romeyka is spoken in a mountainous region at the southern Black Sea coast historically known as Pontus. The mountain range of the Pontic Alps reaches some 4,000 metres in altitude, and the rugged and poorly accessible terrain is possibly a factor in determining regional variation and preservation of archaisms (Özkan 2013). Due to the semi-pastoral lifestyle in the villages and on the mountain pastures (Tr. *yayla*), a rich lifestyle-specific vocabulary in Romeyka remains, including terms for botany, work practices, and products (Schreiber and Sitaridou 2017).

Romeyka arguably consists of three sub-dialects which are mutually intelligible but differ in terms of phonology and lexicon as well as according to some morphosyntactic features (see Mackridge 1987). However, scholars do not agree on the number and grouping of sub-dialects. Figure 1 presents three major sub-dialectal areas: Romeyka of Of (in previous research also referred to as “Ophitic”, henceforth ROf; it is not clear whether Romeyka is still spoken in Of nowadays, cf. Sitaridou 2014b: 32), including Romeyka of Çaykara; Romeyka of Sürmene (henceforth RSür); and Romeyka of Tonya. Brendemoen (2002: 32) identifies another sub-dialect around Maçka which he reports to be due to immigration from

the Sürmene area similar to that of RSür. ROf is argued to be the most archaic variety of Romeyka (Mackridge 1987).

The Romeyka varieties spoken in Pontus in turn are more conservative than Pontic Greek (henceforth PG) due to contact of the latter with Standard Modern Greek (henceforth SMG) after 1932, but also due to the fact that Christian Pontic Greek in Pontus is assumed to have been in contact with other modern Greek varieties prior to 1923.

1.5. Language contact

Romeyka has been in contact with Turkish probably at least since the Islamization waves in the 17th century (Brendemoen 2002) but especially since internal migration started in the 1960s. Arabic elements were introduced to Romeyka via Ottoman Turkish. Intense language contact between Romeyka and Turkish (see Mackridge 1987; Brendemoen 2006) resulted in both lexical and grammatical borrowing and syntactic calquing (Sitaridou 2013).⁴ Romeyka may also have been in contact with Armenian Hemshin/Homshetsma, Georgian, and Laz which is suspected of having exerted a substrate influence on Romeyka (Brendemoen 2002; Sitaridou 2013). Furthermore, Kurdish guest workers must have stayed in the area at several points in history (Sitaridou 2013) though it is not clear whether this had any influence.

1.6. Previous studies and methodological background

There are some early grammar(-fragments), focussing on Pontic Greek in Pontus, in particular Christian Pontic Greek (Deffner 1878; Parcharidis 1880, 1888; Dawkins 1931, 1937;). As Romeyka is expected to have diverged from Christian Pontic Greek already before 1923 (Sitaridou 2014b), these descriptions cannot be uncritically adopted as a baseline for understanding the historical development of Romeyka. The last comprehensive grammar of PG was presented by Drettas (1997). We still lack a thorough typological description of many domains of the grammar of Romeyka although previous publications aimed at summarizing the features of Romeyka that differ from SMG or other PG dialects (Mackridge 1987; Bortone 2009; Özkan 2013). During the last years, a considerable body of research has been carried out on the syntax of Romeyka (see Sitaridou 2013, 2014a, 2014b, and the British Academy Project mentioned in Fn. 1).

The purpose of this chapter is to summarize the information available on Romeyka grammar rather than presenting new data. Apart from above mentioned

⁴ For an investigation of mutual influences of Pontic Greek in Pontus and regional Turkish varieties, especially with regard to the phonological domain, see Brendemoen (2002).

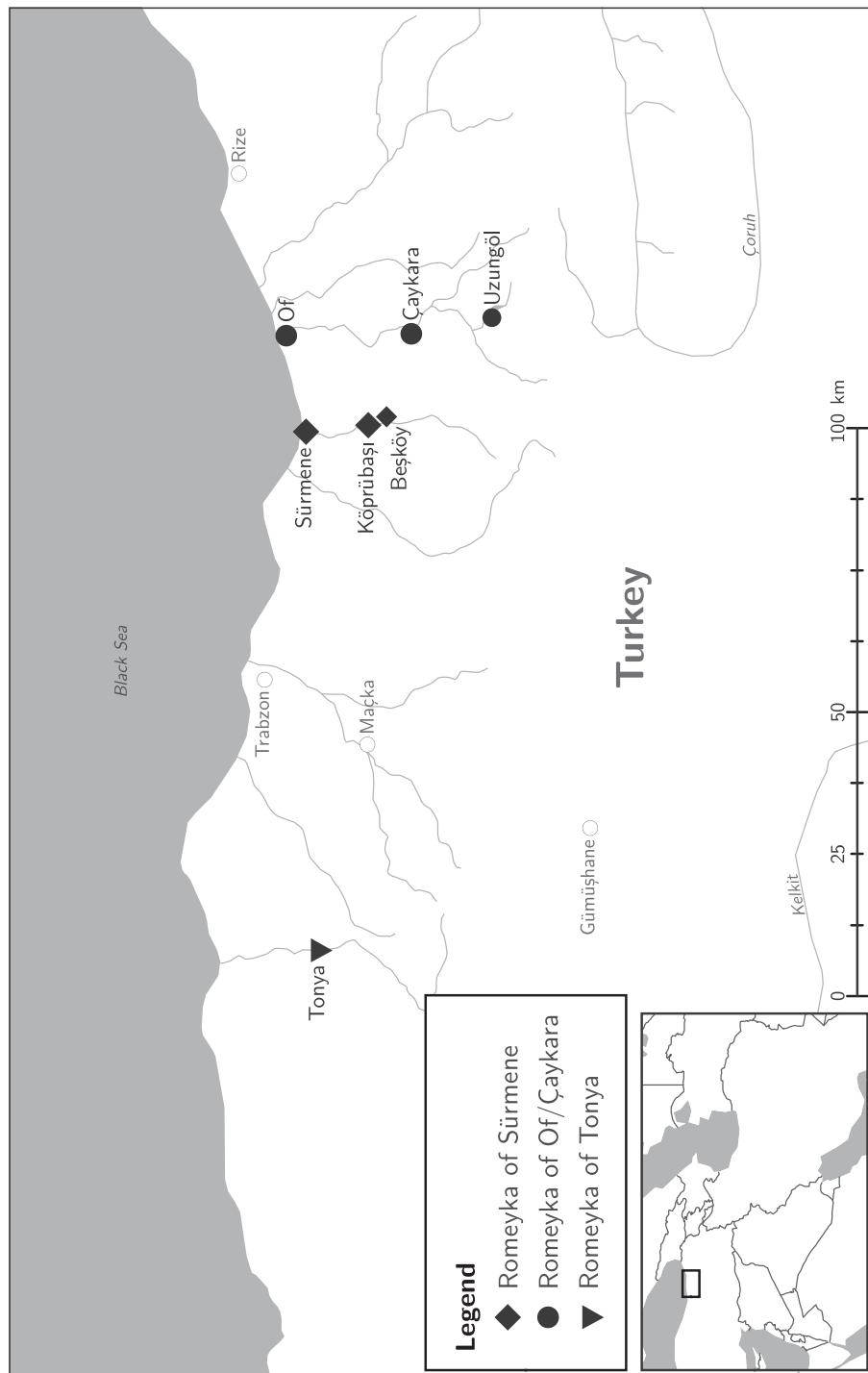


Figure 1: Micro-variation in Romeyka

earlier research, the information in this chapter draw on an unpublished grammar sketch of Özkan (henceforth cited as Özkan, n.d.) as well as on own field work with a female 81-year old Istanbulite L1 speaker of ROF as spoken in Çaykara. The following description does not focus on a particular variety of Romeyka but rather aims to provide general information about particular linguistic features. When information on sub-dialectal variation is available, differences between the varieties are pointed out as such.⁵ The following sections give information on phonology (§2); nominal inflection (§3); determiners, pronouns, and quantifiers (§4); verbal morphology (§5); and syntax (§6) including the syntax of the noun phrase (§6.1), prepositional phrases (§6.2), negation and modality (§6.3), word order in independent (§6.4) and subordinate clauses (§6.5), adverbial (§6.7) and conditional clauses (§6.8).

2. Phonology

To date, no comprehensive analysis of Romeyka phonology is available, nor is there a survey of Romeyka-internal phonological variation. The following notes are gleaned from several sources, and my own fieldwork (see §1.6). Some aspects, in particular those diachronically divergent from SMG or other Greek varieties, have been described by Mackridge (1987, 1999), Bortone (2009), and Özkan (2013). The phonology of Christian Pontic Greek is described by Oikonomidis (1908). Drettas' (1997) grammar of Pontic Greek includes a chapter on phonology. Furthermore, Brendemoen (2002) describes the phonology of the Turkish dialects of Trabzon with reference to interference with Romeyka.

2.1. Segments

2.1.1. Vowels

Romeyka exhibits, in addition to the five vowels also present in SMG, /a/, /e/, /i/, /o/, /u/, also the vowels /æ/ and, more rarely, /œ/ (Mackridge 1987; Özkan 2013). The sub-dialect ROF as spoken in Uzungöl (Sarachos) seems to be the only variety lacking the latter two vowels (Mackridge 1987: 131). According to Mackridge (1987: 121), the vowels /u/ and /y/ are used in Turkish loans, whereby Özkan (2013: 142) points out that these vowels often figure differently in the Trabzon-Turkish dialect, possibly due to contact with Greek (Brendemoen 2002).

⁵ Given the fact that Romeyka lacks any orthography and that its phonemic inventory has not been adequately described yet (Mackridge 1987), Romeyka examples are presented in a provisional broad phonemic transcription; word accent is not indicated.

Unlike in SMG and other Greek varieties, initial unstressed vowels are retained in Romeyka verbs, e. g. *epero* ‘take.1SG’, SMG *παιρνω*, and nouns, see (1a) and (1c) (Mackridge 1987; Bortone 2009). In some varieties of the language, also post-tonic /i/ and /u/ are retained whereas they are dropped in others (1a–b, see also §2.2) (Mackridge 1987; Özkan 2013).

- (1) a. /ospiti/ ROF as spoken in Uzungöl (Sarachos)
 b. /ospit/ RSür as spoken in Beşköy
 c. *σπίτι*, /spiti/ SMG
 ‘house’

In ROF, vowels are lengthened at the end of words and, particular when they are phrase final (Özkan 2013: 139). At present I assume that there is no phonemic length distinction, but more detailed analysis would be necessary to demonstrate this assumption conclusively.

2.1.2. Consonants

The consonantal system of Romeyka is presented in Table 1.

Table 1: The consonantal inventory of Romeyka

		Labial	Dental	Alveolar	Post- alveolar or Palatal	Velar	Uvular	Glottal
Plosive	Voiceless	p		t		k		
	Voiced	b		d		g		
Affricate	Voiceless			ts	tʃ			
	Voiced				dʒ			
Fricative	Voiceless	f	θ	s	ʃ	x		h
	Voiced	v	ð	z	ʒ ⁶	ɣ		
Nasal		m		n				
Tap				r				
Lateral				l				
Approximant					j			

⁶ No example of this phoneme (apart from the consonant cluster /dʒ/) found in the material available to me.

In essence, the consonant system is Greek with few additional phonemes, i. e. in particular the palatal fricatives and affricates which exist in Turkish loans and occur as conditioned variants of palatalized /x/, /g/, and /k/ though not all palatalized velars become palato-alveolars (Mackridge 1987). Özkan (2013: 140–141) reports for RSür as spoken in Beşköy allophonic variation regarding the place of articulation of the voiceless fricatives, e. g. /x/ → /χ/, /x/. Romeyka exhibits at least three affricates /ts/, /tʃ/, and /dʒ/.

The voice distinction is weakened for plosives, with underlying voiceless stops often realized as voiced, though, according to my own fieldwork, that seems to hold true to a lesser extent for the velar plosive /k/. Mackridge (1987: 123) claims for PG that there is, apart from loan words, no distinction of voice in the stops. Stops are generally considered voiced with fortis articulation and become lenis after a nasal (Mackridge 1987; Özkan 2013). The lack of a clear distinction with regard to voicing is reflected in the phonology of the regional Turkish variety (Brendemoen 2002). The velar stops undergo certain phonetic processes: Unlike in SMG, palatalized /k/ before a front vowel becomes /tʃ/ and palatalized /g/ becomes /dʒ/. This process is prevented by a preceding /s/ which becomes /ʃ/, like in *škilos* ‘dog’ (Mackridge 1999, 1987).

2.2. Syllable structure⁷

The canonical syllable pattern in Romeyka is (C)(C)V(C)(C). A word consists minimally of a single vowel, e. g. the definite feminine article *i*.

All consonants can occur in syllable initial position. Unlike SMG, Romeyka has retained initial unstressed vowels in nouns and verbs (see §2.1.1) (Mackridge 1987: 122).

Romeyka allows consonant clusters consisting of two consonants mostly in syllable initial position. Frequently occurring initial consonant clusters are sequences of fricatives and plosives, fricatives and laterals; some examples are: *tše.ra.to* ‘horn’, *pa.tsi* ‘girl’, *sto.ma* ‘mouth’, *an.dras* ‘husband’, *kse.ro* ‘I know’, *o.pse* ‘night’, *ški.lo* ‘dog’, *fti.ra* ‘louse’, *a.spra* ‘star’, *yli.tši* ‘milk’. Consonant clusters in syllable codas are permitted, e. g. *e.mist* ‘we’. Note that the assumed maximum of two consonants in onset or coda can be maintained by considering certain combinations to be affricates, hence single segments. For example, the /tš/ in *i turtš* ‘the Turks’ can be analysed as an affricate, yielding a CVCC syllable. However, the presence and status of affricates in Greek phonology remains generally a controversial issue.

Romeyka exhibits, among others, the following vowel sequences: /ia/, /io/, /ai/, /ae/, /au/, /oi/, e. g. *malia* ‘hair’, *ðio* ‘two’, *fai* ‘bread, food’, *erlaepsa* ‘I understand’, *auros* ‘man’, *oulo* ‘sun’. It cannot always be decided whether they should

⁷ Note that this section is merely based on a 113-item word list from my own fieldwork.

be considered diphthongs (hence filling a syllable nucleus), or whether they should be considered bisyllabic. The combination /ao/ may arise secondarily through the elision of an intermediate velar fricative; this topic requires research.

The combinations /ia/, /io/ are generally realized bisyllabically. Unlike in SMG, they retain their original bi-syllabic structure and resist the tendency to become /j/ + stressed vowel (Mackridge 1987; Bortone 2009). In plural forms, also a sequence of three vowels may occur, as in *faia* ‘bread.PL’. As in PG, no contracted verbs exist; the vowels of the diphthongs /ei/ and /ao/ are separated by a morphemic boundary as in *e-is* ‘you have’, and are likely to be parted by a glide as in *pa-(y)o* ‘I go’ (cf. Drettas 1997: 219). Vowel clusters of two identical vowels do not occur; the masculine and feminine definite articles *o* and *i* are dropped before vowels (Mackridge 1987).

There is significant sub-dialectal variation in the phonological shape of nouns that developed out of AG neuters in *-ion*: In SMG the /on/ ending is lost, whereas ROf as spoken in Uzungöl (Sarachos) retains an older form with /in/, e. g. *omatin* ‘eye’, and RSür as spoken in Beşköy drops the final /n/ (Özkan 2013), e. g. *omati* ‘eye’ (Mackridge 1987: 132). Özkan (2013: 142) states that in varieties dropping unstressed word-final vowels, i. e. *omat* ‘eye’, due to this vowel deletion all consonants can appear in word-final position. According to my own data, most nouns end in a vowel, with only rare instances of voiceless obstruents. It is difficult to identify whether final devoicing takes place as there is in general hardly distinction with regard to voicing (§2.1.2).

3. Nominal inflection

Number, gender and case are the main categories of nominal inflection. Nouns were historically assigned to declensional classes, though the analysis of this system (or its remnants) in Romeyka remains incomplete. Within the noun phrase, articles, adjectives, and some numerals agree with the nominal head in gender, case, and number, although there is variation especially with regard to gender.

3.1. Number

The feature of number in Romeyka has the values singular and plural. The plural is formed by adding a plural suffix, originally selected according to (grammatical) gender (see §3.2): neuter nouns inflect with *-æ* or *-a*, e. g. *ta raši-a* ‘the mountains’; and feminines and masculines by *-ðæs* or *-ðes*, e. g. *i patsi-ðæs* ‘the girls’ (see Table 2 for the paradigm).

Table 2: Number paradigm of RSür as spoken in Beşköy (based on Özkan 2013: 143–144)

	SG	PL	
M	<i>o komšis</i>	<i>i komšiðæs</i>	‘the neighbour(s)’
F	<i>i mana</i>	<i>i manaðæs/manas</i>	‘the mother(s)’
N	<i>t=osbi</i>	<i>ta osbidæ</i>	‘the house(s)’

An older masculine ending *-and* (Mackridge 1987; Drettas 1999) seems to have been superseded by the *-ðæs* ending, as evident in a double set of plural forms, e. g. *babugand/babugaðæs* ‘grandfathers’ (Özkan 2013: 144). The plural in *-and* is only used for some nouns often denoting peoples, families, relatives, professions, and some animals (Özkan 2013: 144), e. g. *i turkand* ‘the Turks’ but cf. *i turtš* ‘the Turks’ in Deffner (1878: 212). A /k/ in the word-final syllable of masculines or feminines may yield a /tš/ plural ending – at least in some varieties, e. g. *turtš* ‘Turks’; *ineka – inetš* ‘women’, presumably reflecting a palatal quality of an underlying plural marker, perhaps *-s*. Another plural suffix is *-ini*, e. g. *aurini* ‘men’.

The original declension classes have increasingly become difficult to identify (Drettas 1997) because of the following factors: (i) masculines and feminines have often come to share the same endings (Özkan 2013); (ii) there has been a tendency to assign masculine and feminine endings only to animate nouns, whereas all inanimate nouns are, irrespective of their grammatical gender, treated as neuter (Dawkins 1931; Drettas 1997; see also Karatsareas 2011 for neuterisation in Cappadocian Greek); (iii) a more general spread of plural neuter forms to masculine and feminine nouns; (iv) some nouns have two plural declensions, e. g. *mana* ‘mother’ in (Table 2). Janse (2002) mentions a possible areal influence of the neighbouring languages not exhibiting gender distinction on the instability of Romeyka nominal inflection.

3.2. Gender

For agreement purposes, Romeyka distinguishes three genders: masculine, feminine, and neuter. Originally, gender was assigned morphologically but shifted toward a more semantically-oriented assignment based on animacy (see Karatsareas 2014; see also §4.1.1). This development, starting in Ancient Greek (henceforth AG) and going further in Romeyka than, for example, in SMG, includes the spread of neuter forms to masculine and feminine paradigms in both singular and plural declension of nouns, determiners, and adjectives (2) (Mackridge 1987; Özkan 2013).

- (2) (Özkan 2013: 144; RSür as spoken in Beşköy)
o vuðias.M.SG – ta vuðia.N.PL ‘the ox(en)’

The spread of neuter applies especially to the plural of [−HUM] feminine nouns and inanimate masculines, and it may even extend to [+HUM] masculine/feminine

nouns. Female proper nouns may be assigned the singular neuter article *to* (Özkan 2013: 145).

In [–HUM] feminine nouns, also mixed declensions occur consisting of a female determiner and neuter adjective (Mackridge 1987).

Gender inflection is also a means to integrate Turkish loans: [+HUM] masculine nouns add *-is* (-*s* following a vowel), while [–HUM] nouns add the neuter suffix *-i* (or *-in* following a consonant), and inanimate nouns ending in a vowel are treated as feminines without suffix, e. g. *i para* ‘money’ (Tr. *para* ‘money’) (Mackridge 1987).

3.3. Case

Romeyka has morphological exponence of nominative, accusative, and genitive cases, expressed on articles and nouns, and other NP constituents. Note that morphological case distinctions on nouns seem to be entirely neutralized in the plural, yielding uniform case-neutral plural endings (cf. Table 3). With the articles, case distinctions are partially retained in the plural, but there seems to be syncretism of plural accusative and plural genitive forms (see Table 4), though there is dialectal variation here. The accusative is used for both the direct and the indirect object, unlike in other Greek dialects where the genitive is used to mark the indirect object (Mackridge 1987). The genitive expresses nominal determination and possession (Drettas 1999). It is not yet known how AG declension classes figure in Romeyka case inflection. Furthermore, there appears to be considerable variation, e. g. *ta peðiði.GEN.PL* ‘boy.GEN.PL’ instead of *ta peðia.GEN.PL* (Özkan 2013).

Table 3: Nominal declension of RSür as spoken in Beşköy (based on Özkan 2013: 143–144)

		SG	PL	
M	NOM	<i>o komši-s</i>	<i>i komši-ðæs</i>	‘the neighbour’
	ACC	<i>ton komši⁸</i>	<i>tsi komši-ðæs</i>	
	GEN	<i>tu⁹ komši</i>	<i>tsi komši-ðæs</i>	
F	NOM	<i>i mana</i>	<i>i mana-ðæs</i>	‘the mother’
	ACC	<i>ti mana</i>	<i>tsi mana-ðæs</i>	
	GEN	<i>tsi mana-s</i>	<i>tsi mana-ðæs</i>	
N	NOM	<i>t[o]=osbi</i>	<i>ta osbid-æ</i>	‘the house’
	ACC	<i>t[o]=osbi</i>	<i>ta osbid-æ</i>	
	GEN	<i>t[u⁹]=osbit-i</i>	<i>ta osbid-æ</i>	

⁸ A word-final /n/ originally occurring in the accusative singular of all genders (in nouns of certain declension classes) is subject to sub-variation and may appear in some nouns, e. g. *ton tširin* ‘the father.ACC’ (Mackridge 1987: 124).

⁹ Only in ROf as spoken in Sarachos (Mackridge 1999) and RSür (Dawkins 1937), other varieties use *tsi*.

The declension of some masculine nominative singular nouns (those of the second declension; in any case only animates) is sensitive to definiteness: the original *-os* ending becomes *-o(n)* when the noun is definite, e. g. *škilos*, *o škilo* '(the) dog' (Mackridge 1987: 124; Dawkins 1931: 394, also for other Asia Minor Greek varieties). According to Dawkins (1931), this phenomenon is caused by the merger of the second and third declension class whereby the nominative of the second declension in *-os* is used for indefinite and the nominative of the third declension in *-o(n)* for definite nouns. Note that the *-o(n)* nominative used for definite nouns coincides with the accusative declension (Dawkins 1937: 31).

Proper nouns denoting male persons are treated like masculine singular nouns: they take the masculine definite article *o*, and the *-is* ending for nominative, e. g. *o Mehmet-is* M.SG.NOM, although there may be a contact-induced tendency to drop the article in the nominative in clause-initial position (for an example for the lack of the article on a clause-initial subject NP see ex. 23 below); this requires more research. Note, however, that Janse (2008: 23) claims for Cappadocian that the nominative became associated with indefiniteness due to syncretism with indefinite animate object NPs and that therefore even definite masc./fem. subject NPs do not take a definite article.

According to Table 3, we would expect a proper noun in direct object function to drop the nominative *-s*. However, in (3) the ending is also dropped for a proper noun in subject function. This suggests that the nominative form only occurs if the constituent concerned is subject, and in sentence-initial position (cf. ex. 41), though this requires further confirmation (cf., for example, Janse (2004: 14) who claims for Cappadocian that due to syncretism of the nominative and indefinite accusative of masc. animate nouns the nominative marker *-s* has been reanalysed as an indefiniteness marker).

- (3) (Michelioudakis and Sitaridou 2012b: 219; ROF, glosses modified)
Eđotšen=eme(n) o Mehmet-Ø ato(n).
 gave.3SG=me.ACC.CL the Mehmet-NOM him/it.ACC
 'Mehmet gave me this/it.'

4. Determiners, pronouns and quantifiers

4.1. Definite and indefinite articles

4.1.1. Definite articles

Romeyka exhibits three definite articles: the masculine singular article *o*, the feminine singular article *i*, and the neuter singular article *to* (see Table 4). ROF as spoken in Uzungöl retains an ancient shape of the masculine/neuter genitive singular article, *tu*, which is *ti/tsi* in other Pontic varieties (Mackridge 1987). The mas-

culine accusative plural article in ROF as spoken in Uzungöl is *tus* which appears elsewhere in the form *tsi/tsu/ti/tin/tis*. Especially in the masculine and feminine plural the articles in Romeyka present an array of different forms (Özkan 2013).

Table 4: Definite articles in Romeyka

		SG	PL
M	NOM	<i>o</i>	<i>i</i>
	ACC	<i>to(n)</i>	<i>tsi/tus</i>
	GEN	<i>tsi/tu</i>	<i>tsi</i>
F	NOM	<i>i</i>	<i>i</i>
	ACC	<i>ti(n)</i>	<i>tsi</i>
	GEN	<i>tsi</i>	<i>tsi</i>
N	NOM	<i>t(o)</i>	<i>t(a)</i>
	ACC	<i>t(o)</i>	<i>t(a)</i>
	GEN	<i>tsi/tu</i>	<i>t(a)</i>

As regards the structure of a NP, an attributively used adjective preceding the noun requires its own definite article, so the NP contains two articles: before the adjective and the noun (cf. 6) (Mackridge 1987). Masculine and feminine definite articles may be dropped before nouns with syllable-initial vowels.

In general, the definite determiners in Romeyka agree with the head in number, gender, and case. However, the spread of the neuter plural declension to [-HUM] feminine and inanimate masculine nouns (see also §3.2), and the spread of neuter adjective inflection to adjectives qualifying a [-HUM] feminine noun, contribute towards a mixed system (Mackridge 1987: 128). Michelioudakis and Sitaridou (2012a: 366, Fn. 3) note that in Romeyka, we find both an inherited grammatical gender agreement system, and a semantic agreement system, the distribution of the two being conditioned by properties of the head such as animacy and gender. Following Karatsareas (2011), they note that position on the Animacy Hierarchy is relevant, with humans behaving distinctly from non-human nouns (see also Karatsareas 2014). In example (4), we witness the initial article *to* indicating semantic agreement (if neuter is considered to reflect non-humanness), while the article immediately adjacent to the head exhibits syntactic (feminine) agreement. This would also be in line with Corbett's (2006: 234) prediction that with stacked agreement targets, as in (4), the target more distant from the controller will show semantic agreement, if any target does (see also Karatsareas 2014). Singular definite articles immediately preceding their heads always exhibit syntactic agreement in Romeyka (Michelioudakis and Sitaridou 2012a: 366, Fn. 3).

- (4) (Mackridge 1987: 128; PG, glosses mine)
- | | | | |
|---------------|---------|-----------|--------|
| to | kokinon | i | kosdra |
| DEF.ART.N | red.N | DEF.ART.F | hen.F |
| ‘the red hen’ | | | |

4.1.2. Indefinite articles

The indefinite article has the form of the numeral ‘one’ *ena* and inflects for case.

4.2. Pronominal system

4.2.1. Personal pronouns

Romeyka has both strong and weak personal pronouns, the latter being enclitically attached to the verb (see (3) for an example). Enclitic pronouns only occur in non-nominative functions. Strong and weak personal pronouns in Romeyka are presented in Table 5. Note that there is a considerable amount of micro-variation in the form of the pronouns; the ones presented here are from ROf.¹⁰

Table 5: Subject and object pronouns in Romeyka

	Subject pronoun	Strong object pronoun	Weak object pronoun (clitic)
1SG	<i>eyo</i>	<i>emena(n)</i>	- <i>m(e)</i>
2SG	<i>esi</i>	<i>esena(n)</i>	- <i>se</i>
3SG	<i>atos.M;</i> <i>ate.F;</i> <i>ato.N</i>	<i>ato(n(a)).N/M;</i> ¹¹ <i>ate(n(a)).F</i>	- <i>((a)n)æ</i>
1PL	<i>emis(t)</i>	<i>emas(una)</i>	- <i>mas(una)</i>
2PL	<i>esis(t)</i>	<i>esas(una)</i>	- <i>sas(una)</i>
3PL	<i>atin</i>	<i>atinus(a)/atinuna</i>	- <i>ata</i>

Object (accusative) pronouns have emerged from AG enclitic personal pronouns (Özkan 2013: 143) and are either strong or weak whereby, as stated by Michelioudakis and Sitaridou (2012b: 219), unlike in other Greek varieties, the use of strong object pronouns seems to be an unmarked option. The third singular object clitic -*æ* seems to be the only third-person form with neutralized gender (Michelioudakis and Sitaridou 2012: 219) although Özkan (2013: 146) distinguishes in RSür between -*((a)n)a.M/N and -*((a)n)æ.F*. In general, the form of the object clitics*

¹⁰ The data stem from Özkan (n.d.), from Michelioudakis and Sitaridou (2012b), Sitaridou (2014b) as well as from own fieldwork.

¹¹ Also *etšinos* ‘this one’ which is actually a demonstrative pronoun.

varies according to whether they are used in transitive or ditransitive verbs, i. e. as to whether clitic clusters occur (but cf. Michelioudakis and Sitaridou 2012b: 237, who argue that there are no clitic clusters in Romeyka, as they question the clitic nature of first and second person object pronouns, the only ones that occur in clusters).

As Michelioudakis and Sitaridou (2012b: 218) argue, if two object suffixes occur together, (weak) Person-Case-Constraint (PCC)-like restrictions apply, such that the third person clitic *-æ* cannot combine with any other clitic to form a cluster, see (5a). However, the combination of a first- and a second-person pronoun (strong PCC) is, unlike in SMG, acceptable (5b).

- (5) (Michelioudakis and Sitaridou 2012b: 218; RSür, glosses modified)

- a. **O Mehmetis eðotše=m=æ*
the Mehmet gave.3SG=me.ACC.CL=it.ACC.CL
'Mehmet gave it to me.'

- (Michelioudakis and Sitaridou 2012b: 238; RSür, glosses modified)

- b. *Eðiksane=m=ese* / **eðiksane=s=eme*
showed.3PL=me.ACC.CL=you.ACC / showed.3PL=you.ACC.CL=me.ACC
'They showed me to you.'

Furthermore, first- and second-person accusative pronouns cannot be interpreted as direct objects in combination with third-person pronouns irrespective of their order (6).

- (6) (Michelioudakis and Sitaridou 2012b: 238; RSür, ROF)

- Eðiksan(e) æ/aton(a) emenan*
showed.3PL him.ACC.CL/him.ACC me.ACC
'They showed him to me/ *They showed me to him.'

Özkan (2013: 146) notes, though, that the third person neuter pronominal suffix can be combined with the neuter definite articles *to* and *ta.PL* used as direct object pronouns after imperative forms like in *ipe-na-to* 'say it' when referring to something previously introduced to the context. For a more detailed discussion of constraints of clitic stacking in Romeyka see Michelioudakis and Sitaridou (2012b).

4.2.2. Demonstratives

Demonstratives have yet to be researched in any detail for Romeyka. Demonstrative pronouns in Romeyka such as *etšinos* 'this one' inflect in general for number, gender, and case. Özkan (n.d.) assumes that the demonstrative pronoun *avudos* or *afsto* 'this' (cf. Trabzon dialectal *habu* 'this' which has arisen due to Greek influence, Brendemoen 2002) can only be used as anaphoric pronouns referring to a previously introduced antecedent, although deictic use seems to be possible, too (7).

- (7) (Own fieldwork; ROF)

Avudo to saxan espero.
DEM the.ACC plate.ACC take.1SG
‘I take this plate.’

The 3rd person neuter pronoun seems also to be used as a demonstrative (8).

- (8) (Own fieldwork; ROF)

Ato to kitap temo en.
3SG.N DEF.ART.N book 1SG.POSS be.3SG
‘This book is mine.’

If a demonstrative determines a noun phrase, the definite article is obligatory before the noun.

4.2.3. Possessives

Possessive pronouns in Romeyka are either enclitics, or independent possessive pronouns (Table 6), which originate from AG possessives (Mackridge 1987; Bortone 2009; Sitaridou 2014b). The third person possessives derive from demonstratives and therefore inflect (in the singular) in accordance with the gender of the possessor (cf. Drettas 1997).

Table 6: Possessive pronouns and pronominal adjectives in Romeyka

	Clitic pronoun	Pronominal adjectives
1SG	-(<i>i</i>) <i>m</i>	<i>temo(n)</i> , <i>tema.N/F.PL</i>
2SG	-(<i>i</i>) <i>s</i>	<i>teso(n)</i> , <i>tesa.N/F.PL</i>
3SG	-(<i>a</i>) <i>t</i>	<i>tatinu.M/N</i> , <i>tatines.F</i> ¹²
1PL	-(<i>e</i>) <i>mun(a)</i>	<i>temetero(n)</i> , <i>temetera.N/F.PL</i>
2PL	-(<i>e</i>) <i>suna</i>	<i>tesetero(n)</i> , <i>tesetera.N/F.PL</i>
3PL	- <i>atun(a)</i>	<i>tatinuna</i> ¹²

The initial vowel of the clitic pronouns may change according to the phonological shape of the noun they are attached to. Prepositions indicating direction, such as *s* ‘to’, fuse with the possessive pronoun, like in *s=temetero* > *semetero to xorio* ‘to our village’.

¹² Drettas (1997) also mentions for Pontic Greek *tetšinu.M/N*, *tetšines.F* as 3rd person singular and *tetšinon* as 3rd person plural possessives for distant possessors.

Another third-person singular possessive suffix used for inanimate objects, animals, and babies is *-(e)θe*, which can apparently be combined with the other possessive clitics (9) (see also Özkan 2013: 146).

- (9) (Sitaridou 2014b: 52; ROf, glosses modified)

<i>Afti</i>	<i>i</i>	<i>ðulian¹³</i>	<i>to</i>	<i>maðin=emun-eθe</i>	<i>yola</i>	<i>en</i>
this.NOM	the.NOM	job.NOM	the	learning=our-its	easy	be.3SG

'It is easy for us to learn how to do this job.'

Possessed nouns require a definite article, regardless of whether the possessive pronoun is independent or clitic (10a–b). Independent possessive pronouns are generally preposed (10a), as are possessive nouns (11a). A possessed NP can itself occur recursively as a preposed possessor, as in (11b). A kind of postponed possessor seems also to be possible, as shown in (12).

- (10) (Own fieldwork; ROf)

- a. *temetero* *t=ospī*
1SG.POSS the=house
- b. *t=ospī=temona*
the=house=1PL.POSS
'our house'

- (11) (Own fieldwork; ROf)

- a. *tsi* *Aišes* *to* *ketši*
the.GEN Ayse.GEN the.NOM goat.NOM
'Ayse's goat'
- b. *tsi* *patsis=im* *to* *ketši*
the.GEN daughter.GEN=1SG.POSS the.NOM goat.NOM
'my daughter's goat'

- (12) (Own fieldwork; ROf)

- t=ospī* *tin* *ates*
the.NOM=house.NOM the.ACC(?) she.GEN
'her house'

4.2.4. Relative pronouns and relativizers

Relative pronouns, or more generally relativizers, are discussed in §6.5.1.1 below. A possible candidate for a relative pronoun is a form derived from the neuter article, *d(o)/d=/t=*.¹⁴ It seems to be used with inanimate, and animate non-human

¹³ Janse (2002: 221) notes that in some cases the possessor of a genitive phrase occurs without overt genitive marking, presumably modelled on the Turkish compound-type of possessive construction.

¹⁴ Özkan (2013) reports the form *nto* in RSür.

nouns (Özkan, n.d.) and for the relativization of objects (13), spatial complements, subjects, and as a free relativizer. Note that this kind of relativization is post-nominal, as opposed to the dominant prenominal strategy discussed in §6.5.1.1 below.

- (13) (Gandon 2016: 221; Romeyka, glosses modified)

Eyrapses ēfinon do ipamen
record.PST.2SG that.ACC REL say.PST.1PL
'Did you record what we have said?'

Variable relativizers are those derived from the interrogative pronouns, e. g. *opios* and *otinan* only used for animates/humans (Özkan, n.d.). Furthermore, Sitaridou (2014b: 30) notes the use of *itina*, an ancient relativizer.

4.2.5. Interrogatives

The information in this section stem from Michelioudakis and Sitaridou (2012a) on *wh*-fronting in PG and Romeyka. *Wh*-words in Romeyka are presented in Table 7. Sitaridou (2014b: 30) mentions furthermore retention of AG *hoθen* < *ōθev* 'wherefore'.

Table 7: Interrogatives in Romeyka (adapted from Michelioudakis and Sitaridou 2012a: 356–358)

	+HUM		−HUM	
NOM	<i>pios</i>	'who'	<i>pios</i> ¹⁵	'what'
	<i>kaškiši</i>	'how many'	<i>pion</i>	'what'
ACC	<i>tinan</i>	'who'	<i>pion</i>	'what'
	<i>motinan</i>	'with whom'	<i>do</i>	'what'
	<i>kaškišus</i>	'how many'	<i>doxna(n)</i> ¹⁶	'what'
Other	<i>layo</i>	'how'		
	<i>pote</i>	'when'		
	<i>pu merea/tšeka</i>	'where.side/there'		
	<i>apoxen</i>	'from.where'		
	<i>poso</i>	'how much'		
	<i>ya to pio</i>	'why'		

According to Michelioudakis and Sitaridou (2012a: 357), interrogatives in Romeyka have grammaticalised [\pm HUM] restrictions but lack number/gender distinctions. The Turkish loan *kaš(i)kiši* 'how many' is optionally used together with *pios* 'who' to add an indication of plurality (14).

¹⁵ Note that this source is difficult to interpret at this point.

¹⁶ In ROF only (Michelioudakis and Sitaridou 2012a: 358).

- (14) (Michelioudakis and Sitaridou 2012a: 358; ROF)

Pios kaš(i)kisi erθen?
 Who NOM.PL came.3SG
 ‘Who (=many) came?’

All *wh*-phrases appear obligatorily in the left periphery, where two or more *wh*-phrases can be fronted (Michelioudakis and Sitaridou 2012a: 360). Michelioudakis and Sitaridou (2012a: 363) argue that movement is obligatory as *wh*-elements are not permitted *in situ*. Apart from discourse-linked interrogatives, multiple *wh*-fronting is order-preserving (15).

- (15) (Michelioudakis and Sitaridou 2012a: 363; ROF)

Pios tinan pote efilise?
 who.NOM who.ACC.HUM when kissed.3SG
 ‘Who kissed whom and when?’

Romeyka has borrowed the Turkish interrogative particle *ml*, used in polar questions, albeit without vowel harmony (16). Like in Turkish, the question particle is flexible to move for focalization purposes. Note that in this example, the *wh*-word *layo* is *in situ*; whether this is evidence of syntactic influence from Turkish, or has other causes, cannot be discerned.

- (16) (Own fieldwork; ROF)

Teso i patsi layo en, emorfo mi en?
 2SG.POSS the.NOM daughter.NOM how be.3SG beautiful Q be.3SG
 ‘How is your daughter, is she beautiful?’

4.2.6. Quantifiers and numerals

Quantifiers are at least *olos* (17) and *ul(os)* ‘all’ (18). The status of the following is unclear: *kat* ‘some’, *kanis* ‘nobody’, *tipu* ‘nothing’ (Özkan, n.d.). At least the first two can take personal clitics (19).

- (17) (Own fieldwork; ROF)

Olon emorfo ospi temo en.
 all.ACC beautiful house 1SG.POSS be.3SG
 ‘The most beautiful house is mine.’

- (18) (Own fieldwork; ROF)

So Istanbul op= ine ul Rumdža u= kserone.
 in.the Istanbul REL= be.3PL all.NOM Rumca NEG= know.3PL
 ‘The people (who are) in Istanbul all don’t know Rumca.’

- (19) (Own fieldwork; ROF)

Eyo ul=isuna eyabo.
1SG all=2PL.ACC love.1SG
'I love you all.'

Romeyka numerals exist only for the numbers 'one' to 'five', the rest is Turkish (Mackridge 1987). The numeral *ena* 'one' can occur determined, e. g. *to ena* 'the one'. Furthermore, pronominal clitics can be attached to numerals, though the personal suffix in (20) resembles the Turkish first plural suffix *-Iz*. Similarly, Tr. *tek* 'one, only' may be used emphatically, e. g. *ena tek yarðel* 'just one child', cf. Turkish *bir tek çocuk*.

- (20) (Own fieldwork; ROF)

na troyume i ði-jiz
FUT eat.1PL the two-1PL
'The two of us will eat.'

5. Verbal morphology

Romeyka finite verbs encode person, TAM, and voice. Some functions are expressed peripherastically by the use of particles, such as the future marking particle *na*, and others by enclitic affixes and/or stem alternation. The verbal system of Romeyka contains many archaisms absent in other Modern Greek dialects such as the AG aorist infinitive (see Mackridge 1987; Sitaridou 2014b).

5.1. Person agreement

Romeyka verbs are specified for person and number, realized through personal suffixes agreeing with the subject. The clitic pronouns for objects (cf. Table 5 above) form a distinct paradigm; whether they should be considered "agreement" is an issue we do not take up here. In what follows, I refer exclusively to the subject-indexing suffixes as agreement. The paradigms vary according to verb stems. Some verbs are strong, i. e. they exhibit different stems with regard to tense.

The citation form used here is the first person singular present. Verbs can be classified according to the endings in this form: verbs ending in *-o*, e. g. *tro(y)o* 'eat', and verbs ending in *-me*, e. g. *tšimume* 'sleep', (Table 8 for the full paradigm based on own fieldwork). Within these classes, verbs can be subcategorized according to ultimate and penultimate stress, the latter verb class forming the major group (Drettas 1997: 205).

Table 8: Present tense inflectional paradigm of Romeyka

	Verbs in <i>-o</i>	Verbs in <i>-me</i>
	'eat'	'sleep'
1SG	<i>tro-(y)o</i>	<i>tšim-ume</i>
2SG	<i>tro-s</i>	<i>tšim-ase</i>
3SG	<i>tro-i</i>	<i>tšim-ate</i>
1PL	<i>tro-yume</i>	<i>tšim-umist</i>
2PL	<i>tro-te</i>	<i>tšim-asten</i>
3PL	<i>tro-yune</i>	<i>tšim-untane</i>

Strong verbs have stems differing in phonologically unpredictable ways, or through suppletion. They can be divided into two groups: (I) those with an imperfective stem, used for e. g. the present tense, and a deviating perfective (traditionally aorist) stem, e. g. *elepo* 'see.PRS', *elepa* 'see.IPFV', *iða* 'see.AOR; (II) those with a present stem different from the aorist and imperfect stem, e. g. *eftao* 'do.PRS', *epina* 'do.IPFV', *epika* 'do.AOR' (Drettas 1997: 205).

5.2. Tense

Romeyka has three morphologically marked tenses: present, imperfect, and aorist. Other tenses such as the future are formed peripherastically. Remarkably, there is no periphrastic perfect tense, e. g. with 'have' (Mackridge 1987: 127).¹⁷ Note, however, periphrastic present progressives which are discussed in §5.3.

5.2.1. Present

The present tense expresses an imperfective aspect and is used for habitual actions (21). It is formed on the basis of the imperfective stem. For the paradigm see Table 8.

- (21) (Own fieldwork; ROF)

Her sabaxtan so mektep payo
 every morning to.the school go.1sg
 'Every morning I go to school.'

¹⁷ See Sitaridou (2014a: 122) for the discussion of a complex predicate consisting of the verb 'have' in past tense followed by an infinitive which resembles a past perfect though only functioning as a counterfactual.

5.2.2. Imperfect

The imperfect expresses an action which happened continuously or habitually in the past. Like the aorist, it is formed by the vocalic augment /e/ (Bortone 2009), like in *emaireva* ‘I used to cook.IPFV’ vs. *emairepsa* ‘I cooked.AOR’. As in the other tenses, the inflectional paradigm depends on the verb class (cf. Drettas 1997 for PG). In the speech of the Istanbulite ROF consultant, however, no imperfect forms were attested. Instead, the speaker used an aorist in the relevant context, as in (22):

- (22) (Own fieldwork; ROF)

Mikrina anda=emunest-ten emist-pal so mektep epixame
 small(?) time.in=our? we-FOC to.the school go.AOR.1PL
 ‘When we were small, we used to go to school.’

5.2.3. Aorist

The aorist is used for perfective actions in the past (23). It can be combined with temporal adverbs or inchoative verbs to mark the beginning of an action carrying on to the present. Furthermore, it can be used (together with the future particle *na*) for imminent actions happening in the near future (24). It is formed on the basis of perfective aspect, i. e. the aorist stem with the ancient temporal augment /e/ (Table 9), its forms matching largely with the AG aorist endings (Sitaridou 2014b: 53).

- (23) (Own fieldwork; ROF)

oretmenis aso xorio erθe
 teacher.NOM from.the village come.AOR.3SG
 ‘The teacher came from the village.’

- (24) (Own fieldwork; ROF)

t=akšam n=arde
 the=evening FUT=come.AOR.3SG
 ‘She will come this evening.’

Table 9: Aorist endings in ROF

	Verbs in <i>-o</i>	Verbs in <i>-me</i>
	'ate'	'slept'
1SG	<i>e-fa-ya</i>	<i>e-tšim-eθa</i>
2SG	<i>e-fa-es</i> ¹⁸	<i>e-tšim-eθes</i>
3SG	<i>e-fa-e</i>	<i>e-tšim-eθi</i>
1PL	<i>e-fa-yame</i>	<i>e-tšim-eθame</i>
2PL	<i>e-fa-ete</i>	<i>e-tšim-eθete</i>
3PL	<i>e-fa-yane</i>	<i>e-tšim-eθane</i>

5.2.4. Future

There is no morphological future tense in Romeyka. Future is expressed periphrastically by means of the particle *na* (25a) and the present stem, like in Medieval Greek and unlike in SMG, which uses the particle *θa*. *Na* is, apart from being a future marker, also used as a subordinating conjunction after volitives and other verbs (Sitaridou 2014a; see §6.5.1.3). We nevertheless gloss it with FUT throughout, although this is not always an accurate reflection of the function. For RSür as spoken in Beşköy, Özkan (2013: 147) also mentions a preverbal particle *ha* used to form future in negated sentences (cf. 25b) and interrogatives.

(25) (Own fieldwork; ROF)

- a. *Sabahle na pame so dyyun.*
tomorrow FUT go.1PL to.the wedding
'Tomorrow we will go to the wedding.'
- b. *Utše na pa(y)o*
not FUT go.1SG
'I will not go.'

5.3. Aspect

Perfective and imperfective aspect are in Romeyka only realized in the past tense indicative, i. e. by the aorist and the imperfect respectively (Mackridge 1987; Sitaridou 2014a). Present and future have no morphological aspectual distinctions (cf. 26).

(26) (Özkan, n.d.; RSür as spoken in Beşköy, glosses mine)

- t=ospı tamır na inete*
the=house repair FUT become.3SG
'the house will have been repaired'

¹⁸ Not attested in my data.

Progressive aspect seems to be realized by two strategies: (i) by the invariable auxiliary *eš(i)* + a finite verb form of the present stem for present (27a) and in the imperfect for past (Özkan 2013). Istanbulite ROF applies the aorist form instead (27b). Invariable *eš(i)* resembles the 3SG form of the verb ‘to have’ *eš*, and could therefore be a frozen grammaticalized form of ‘have’ (see also Drettas 1997: 334). Note the use of ‘have’ as an auxiliary for progressive aspect in vernacular Persian (see Paul, this volume, chapter 4.6).

(27) (Own fieldwork; ROF)

- a. *Aso pazar eš erxume.*
from.the market AUX come.PRS.1SG
'I am coming from the market.'
- b. *I Ayše aso Istanbul eš erte.*
the Ayse from.the Istanbul AUX come.AOR.3SG
'Ayse has been coming from Istanbul.'

In PG, *eš + k* ‘and’ (*tše* in Romeyka) + present/imperfect is used to express explicit processes which are near to completion (Drettas 1997).

(ii) Another periphrastic construction consists of a finite verb + coordinative *tše* ‘and’ + finite *steko/stekume* ‘stand’ or *kahome* ‘sit’. Both auxiliaries seem to be used interchangeably whereby *steko* and *stekume* seem to constitute a mixed paradigm in themselves (see Table 10).

Table 10: Periphrastic progressive forms with *kahome* ‘sit’ and *steko/stekume* ‘stand’ on the example of *tšimume* ‘sleep’¹⁹

<i>kahome</i> ‘sit’		<i>steko/stekume</i> ‘stand’
'sleep' + 'and' + 'sit'		'sleep' + 'and' + 'stand'
1SG	<i>tšimume tše kahome</i>	<i>tšimume tše steko/stekume</i>
2SG	<i>tšimase tše kahese</i>	?
3SG	<i>tšimate tše kahete</i>	<i>tšimate tše stitš</i>
1PL	<i>tšimumist tše kahomist</i>	?
2PL	<i>tšimaste tše kahesten</i>	?
3PL	<i>tšimun tše kahontane</i>	<i>tšimun tše stekontane/stekune</i>

In the past, both verbs are in the imperfective stem (Özkan 2013); cf. Istanbulite ROF in (28) which uses the aorist with the present form of the auxiliary *steko*.

According to Özkan (2013: 148), the construction is used especially for “longer lasting” progressive actions (see ex. 29) and possibly replicates a Turkish periphrastic progressive -*Ip + dur-*, like in *çalışıp duruyorum* ‘I keep working’ (see

¹⁹ For the forms of *tše steko/stekume* marked by a question mark no information was available in my data.

also Dawkins 1916: 199 who reports the same construction with ‘stop’ and ‘lie’ in other AMG varieties).

- (28) (Own fieldwork; ROF)

<i>to</i>	<i>vakit</i>	<i>erθe</i>	<i>(tše stitš)</i>	<i>n=efta</i>
the	time	come.AOR.3SG	(and stand.PRS.1SG)	FUT=make.1SG
<i>to</i>	<i>faji</i>			
the	food			

‘The time has come, I will prepare the food.’

- (29) (Own fieldwork; ROF)

<i>So</i>	<i>mutfak</i>	<i>tšališevo</i>	<i>tše</i>	<i>steko.</i>
in.the	kitchen	work.1SG	and	stand.1SG

‘I am working in the kitchen.’

According to Sitaridou (2014b: 44), the construction with *steko* bears inchoative aspect though without a progressive function (cf. Drettas 1997: 336).

5.3.1. Voice

The details of passive morphology in contemporary Romeyka are yet to be established. Historically, the categorization of verbs in two classes approximately corresponds to a voice distinction: verbs in *-o* are called actives whereas verbs in *-me* are called (medio)passives (Drettas 1997: 205). Whereas ROF seems to lack passives in general, RSür allows passives (Michelioudakis and Sitaridou 2012b: 236). The syntax of datives under passivization in RSür has been investigated by Michelioudakis and Sitaridou (2012b). They note that a theme argument can be a regular subject of a passive verb (30–31a), while a benefactive or recipient cannot advance to subject under passivization (31b):

- (30) (Michelioudakis and Sitaridou 2012b: 236; RSür, glosses modified)

<i>To</i>	<i>harti</i>	<i>eyrafte</i>	<i>tin</i>	<i>Aiše</i>
the.NOM	letter.NOM	written.PASS.3SG	the.ACC	Aise.ACC

‘The letter was written for Aise.’

- (31) (Michelioudakis and Sitaridou 2012b: 236; RSür, glosses modified)

- a. *I* *para* *tin* *Aiše* *eđoste.*
 the.NOM money.NOM the.ACC Aise.ACC given.PASS.3SG
 ‘The money was given to Aise.’
- b. **I* *Aiše* *eđoste* *tin* *paran*
 the.NOM Aise.NOM given.PASS.3SG the.ACC money.ACC
 ‘Aise was given the money.’

5.3.2. Mood

Romeyka exhibits four moods: Indicative (*tros* ‘you eat’) and imperative (*fa* ‘Eat!’) are morphologically marked on the verb whereas subjunctive (*na troyo* ‘I should eat’) and optative (*as troyo* ‘let me eat’) are formed by the particles *na* and *as*, respectively, followed by the finite verb (Sitaridou 2014a).

The imperative covers the second persons and differs according to verb class: verbs in *-o* such as *tro(y)o* ‘eat’, *fa* ‘eat.IMP.2SG’, *fate* ‘eat.IMP.2PL’; verbs in *-me* such as *tšimume* ‘sleep’, *tšimeθ* ‘sleep.IMP.2SG’, *tšimeθisten* ‘sleep.IMP.2PL’ (see Drettas 1997: 227–232). Negation is formed by the negation particle *mi* and – at least in Istanbulite ROF – the present indicative of the second persons, e. g. *mi tros* ‘NEG eat.PRS.2SG’. Thus ROF neutralizes the imperative/indicative distinction under negation. Some verbs in Romeyka retain the ancient imperative ending *(-s)on*, e. g. *akuso(n)* ‘listen’ (Bortone 2009: 84), but also other archaic imperatives such as *ipe* ‘say’ (Mackridge 1987: 125). On the syntax of modality see §6.3.

5.4. Finiteness

Mackridge (1987: 127, Fn. 17) considers that Romeyka lacks non-finite forms of the verb, based on the assumption that finiteness equates to “indexing person”. However, the infinitive found in certain varieties of Romeyka seems to run counter to this (32) (see Sitaridou 2014b and the discussion in §6.5.2.1). Although there appears to be no gerund in Romeyka (Sitaridou 2014a), Özkan (n.d.) provides examples from RSür resembling adverbial past participles which are formed by the suffix *-ta* appended to the perfective stem, e. g. *yelax-ta erθame* ‘we came laughing’.

- (32) (Sitaridou 2014a: 122; ROF, glosses modified)

<i>Uts</i>	<i>eporesa</i>	<i>tšimiθini</i> .
not	can.PST.1SG	sleep.INF
‘I could not sleep.’		

5.5. Verbal derivation

This section is based on Özkan (2013: 147). Verbal derivation is productive in Romeyka. A variety of derivational affixes is used for verbalization, causativization, or integration of Turkish loans. A root stem to which derivational suffixes are attached forms the verbal base on which the inflection suffixes are attached. A single affix can have several functions. For example, the suffix *-(i)zo* is used: (i) for verbalization of nouns, e. g. *fos* ‘light’ → *fosizo* ‘to be lit/to shine’; (ii) to increase transitivity or as a causative marker, e. g. *maθano* ‘to learn’ → *maθizo* ‘to teach’; (iii) for the formation of onomatopoetic verbs such as *tsatsalizo* ‘to weed’; (iv) to grammatically adapt Turkish verbs like in *dajanizo* ‘to endure’ from Tr. *dayan-* ‘to endure’.

This range of functions matches very closely the *and* suffix of Kurmanji Kurdish, discussed in Haig (this volume, chapter 2.3). Another frequent suffix used for integrating Turkish loans is *-evo*, e. g. *tšališevo* ‘to work’ from Tr. *çalış-* ‘to work’.

6. Syntax

6.1. Syntax of the noun phrase

Noun phrases (henceforth NPs) in Romeyka are head-final:

- (33) (Constructed example)

<i>o</i>	<i>emorfos</i>	<i>o</i>	<i>peðas</i>
the.M	handsome.M	the.M	boy.M
‘the handsome boy’			

The definite article co-occurs with all nouns, including proper nouns, and co-occurs with both demonstratives, and possessives (34). Quantifiers are generally treated like adjectives and thus require their own article (35).

- (34) (Own fieldwork; ROf)

<i>tatinuna</i>	<i>to</i>	<i>xorio</i>
3PL	the	village
‘their village’		

- (35) (Own fieldwork; ROf)

<i>T=ulinuna</i>	<i>t=ospidæ</i>	<i>mikra</i>	<i>ine.</i>
the=all	the=houses	small	be.3PL
‘Everybody’s houses are small.’			

Adjectives in Romeyka precede the noun they modify, and generally agree with their heads in person and gender, though there may also be semantic agreement (Michelioudakis and Sitaridou 2012a: 366, Fn. 3). In contrast, Özkan (2013: 145) suggests that syntactic agreement is not consistent with adnominal adjectives (36), while predicatively used adjectives exhibit semantic agreement if the noun denotes a human referent (at least this is how his example can be interpreted). There is obviously dialectal variation here, but the general tendency for the neuter form to expand its range appears to be valid here (cf. Karatsareas 2011, 2014).

- (36) (Özkan 2013: 145; RSür)

<i>do</i>	<i>galo</i>	<i>i</i>	<i>mana</i>
the.N	good.N	the.F	woman.F
‘the good woman’			

Furthermore, a feminine animate noun may be modified by one neuter and one masculine adjective (Mackridge 1987: 128).

6.2. Prepositions and prepositional phrases

Romeyka is a prepositional language. The local prepositions *s* ‘to, in’ and *as* ‘from’ merge with the definite articles; *s > to*, *s=o* ‘in the’. Most of the prepositions require accusative case, though the local prepositions may also govern a genitive-NP (Özkan 2013: 145). Note that in Istanbulite ROF the local prepositions seem not to take accusative case, but rather the nominative. *Me* and *ja* become *met* and *jat* when their complement is a pronoun. For an overview of some prepositions and their function see Table 11.

Table 11: Prepositions in Romeyka

Preposition	Function	Equivalent
<i>os</i>	temporal	until
<i>prin</i>	temporal	before
<i>s</i>	locative	in, to
<i>as</i>	ablative	from, of, by, from ... on, since
<i>ja</i>	purpose	for (RSür)
<i>ðæ</i>	purpose	for (the sake of) (ROf)
<i>me</i>	instrumental	with, through, by

6.3. Negation and modality

This section is based on Sitaridou (2014a) on modality in Romeyka. Romeyka expresses negation by the use of negators with rich allomorphic and cross-dialectal variation, whereby the allomorphic distribution of the negators correspond to that of AG (Sitaridou 2014b). The following four negators can be distinguished: (i) *utš* for sentential negation in indicative sentences (37a), (ii) *mi(n)* in imperatives and wishes (37b), (iii) *xe* in subjunctives (37c), (iv) *mutš* in counterfactuals (Sitaridou 2014a: 121; see Table 12 for an overview).

(37) (Own fieldwork; ROF)

- a. *Utš eyabo na troyo.*
NEG want.1SG FUT eat.1SG
'I don't want to eat.'
- b. *Esist mi trote, kalo utš en.*
2PL NEG eat.IMP.2PL good NEG be.3SG
'Don't eat [that], [it] is not good.'
- (Sitaridou 2014a: 123; ROF, glosses modified)
 - c. *Esi ðelis eyo xe na troyo*
you.NOM want.2SG I.NOM not FUT eat.1SG
'You don't want me to eat.'

Utš is derived from the ancient pre-verbal negative particle *oúk*. Its phonologically conditioned allomorphic forms are presented in Table 12.

Table 12: Negators in Romeyka (adapted from Table 1 in Sitaridou 2014a: 121)

Negator	Conditions	Example
<i>utš</i>	<i>utš</i> phonologically conditioned variation (see Sitaridou 2014a: 121 for details)	<i>utš eyapo</i> ‘I don’t like’
	<i>tš</i>	<i>tš alo</i> ‘not more’
	<i>tši</i>	<i>tši pao</i> ‘I don’t go’ ²⁰
	<i>u</i>	<i>u θelo</i> ‘I don’t want’
	<i>utše</i>	<i>utše na troyo</i> ‘I wouldn’t eat’
<i>mi(n)</i>	imperatives; wishes	<i>mi trote</i> ‘Don’t you eat?’
<i>xe</i>	subjunctives	(see 37c)
<i>mutš</i>	counterfactuals	<i>na mutšixa xasini</i> ‘If I had not lost’ ²¹

Contact with Turkish affected the responsive particles used in ROF as spoken in Uzungöl (Sarachos): whereas Greek *ne* is used for ‘yes’, ‘no’ is expressed by the Turkish *yok* (Mackridge 1987: 134).

The periphrastic subjunctive formed by the particle *na* and the present stem of the verb (or the imperfect in a *na*-clause governed by an aorist verb, Sitaridou 2014a) is used for a variety of constructions, such as volitionals (discussed in §6.5.1.3 on *na*-clauses) and conditionals (discussed in §6.8). There is no aorist subjunctive (but cf. Dawkins 1937 who claims the aorist subjunctive survived in the PG of Christians from Samsun). *Na* as a future marker often also seems to imply a modal meaning (38).

- (38) (Own fieldwork; ROF)

ðo=me nero, na pino.
give.IMP.2SG=1SG water FUT drink.1SG
'Give me water, (that) I may drink.'

Optative mood is formed by the particle *as* and the verb in present tense. It is used in voluntatives (39) and counterfactuals (40).

- (39) (Own fieldwork; ROF)

Elate, as troyume!
come.IMP.2PL OPT eat.1PL
'Come, let us eat.'

²⁰ Sitaridou (2014a: 130).

²¹ Chatzopoulou and Sitaridou (2014, ex. 7).

- (40) (Own fieldwork; ROF)

As eftayo=se faji, ama utše na troyo ipes me.
 OPT make.1SG=2SG food but NEG FUT eat.1SG say.2SG 1SG
 ‘I had made you food, but you said to me you wouldn’t eat it.’

Romeyka allows for embedded imperatives as complements to verbs of saying (Sitaridou 2014a) such as in (41) although it is not clear to what extent this construction differs from indirect speech. The order of the subordinate imperative and the main verb are variable.

- (41) (Sitaridou 2014a: 129; ROF, glosses modified)

O Alis esenan yrapson ipen
 the Alis.NOM you write.IMP.2SG tell.PST.3SG
 ‘Alis ordered you to write.’

6.4. Word order in independent clauses

It remains controversial whether the basic word order in Romeyka is predominantly head-final (42) though permitting VO order (see for example ex. 25a) (Michelioudakis and Sitaridou 2012b: 216) or whether Romeyka’s canonical order is SVO, with OV (see for example ex. 27) resulting mostly from focus fronting and interrogative clauses, as well as from Turkish contact (Neocleous 2017).

- (42) (Michelioudakis and Sitaridou 2012b: 216; ROF)

škilon exo
 dog.ACC have.1SG
 ‘I have a dog.’

6.4.1. Ditransitives and causative constructions

This section is based on the description of ditransitives and benefactives in Romeyka in Michelioudakis and Sitaridou (2012b: 231–236). In ditransitives, recipients are accusative and do not alternate with prepositional phrases (43). Both orders of direct and indirect object seem to be acceptable, though direct object – indirect object order appears to be preferred (see Michelioudakis and Sitaridou 2012b).

- (43) (Michelioudakis and Sitaridou 2012b: 231; RSür, glosses modified)

*To peđi eđotše fai ton ađelfo. / *son ađelfo.*
 the child gave.3SG food the.ACC brother.ACC / *to.the brother
 ‘The child gave food to the brother.’

In benefactives, again both word orders are possible (44). Benefactives may alternate with prepositional phrases headed by *đæ* ‘for’ (ROF) or *ja* ‘for’ (RSür).

- (44) (Micheloudakis and Sitaridou 2012b: 235; RSür, glosses modified)

*Aiše epitše to Mehmet pide /
 Aise made.3SG the.ACC Mehmet.ACC pie.ACC /
 pide to Mehmet.
 pie.ACC the.ACC Mehmet.ACC
 ‘Aise baked Mehmet a pie.’*

Causatives may be expressed peripherastically using a matrix verb ‘let, allow’, and a complement clause introduced by *na* (45) (Sitaridou 2014a: 125). Apart from this, causativity in Romeyka awaits further investigation.

- (45) (Sitaridou 2014a: 125; ROf, glosses modified)

*Efikane=sas na skafete ta xorafæ=suna.
 let.PST.3PL=you FUT dig.2PL the fields=his
 ‘They let you dig his fields.’*

6.5. Subordinate clauses

Romeyka exhibits several different strategies for subordination, including finite and non-finite verbal forms (Sitaridou 2014a, 2014b). The only complementiser in regular use is the all-purpose modal particle *na* (see §6.5.1.3; §5.2.4). In what follows, a broad distinction is drawn between finite and non-finite subordination distinguishing several sub-types.

6.5.1. Finite subordination

6.5.1.1. Relative clauses

There is little information to date on Romeyka relative pronouns (see §4.2.4 above), except Gandon (2016: 219–223). Drettas (1997: 365) notes that relativizers in PG may be analysed as connectors rather than subordinating pronouns. Relativizers in Romeyka are in general preverbal and different strategies of relativization seem to apply according to the role of the head noun of the relative clause. Invariable relativizers are the following, although relative clauses without relativizer are possible, too (see Gandon 2016):

(i) *op* (possibly > AG *όποιος*, Gandon 2016: 221) occurs in prenominal relative clauses, immediately preceding the relative clause verb. It may be used for relativizing the subject of the relative clause, and can be used as a free relative, as in (46).²²

²² Gandon (2016: 223, Fn. 130) notes the possibly nominalizing function of the free relativizers. If analysed as a nominalization, the gloss should indicate NMLZ.ACC.

- (46) (Own fieldwork; ROf)

Aso Katoχor op erθe iða=na
 from.the Katochor REL came.3SG saw.1SG=3SG
 ‘I saw the one that came from Katochor.’

- (ii) Though likely, it is not fully clear whether *p(i)/p(u)* (Özkan), *pe* (> AG *óπερ*, Sitaridou 2014b: 30) derive from the same root as *op*; a clitic *p=* is also found in relative clauses:

- (47) (Gandon 2016: 222, citing Neocleous, in preparation; Romeyka, glosses modified)

Opse iða aλis p=epiren iŋeka
 yesterday saw.1SG Alis.NOM REL=got.married.3SG woman.ACC
 ‘Yesterday I saw the woman who Ali married.’²³

The comparative construction in (48) also exhibits *p=*, though here it is not attached to a verb. This kind of construction suggests that *p=* may also function as a general nominalizer ('that of you'). However, this is uncertain; it may also be analysed as part of the preposition, or it may have contracted with the preposition *aso* 'from' (for an alternative analysis see ex. 69 in §6.7.2).

- (48) (Own fieldwork; ROf)

Temetero t=ospι a=p=esetero trani-dero en
 1PL.POSS the=house.NOM from=REL=2PL.POSS big-CMP be.3SG
 ‘Our house is bigger than yours.’; lit. ‘...than that of you.’

Romeyka uses a typologically uncommon strategy for relativization, involving prenominal, but finite relative clauses (Gandon 2016: 220, citing Neocleous, in preparation) (49b). Right-branching structures (postnominal relative clauses) may be possible, too (49a), presumably a reflex of the type widely-attested in AG, and still maintained in SMG. The prenominal relative clauses (for example exs. 50, 51) in Romeyka are probably contact-induced by Turkish prenominal relative clauses. There are also evident similarities to geographically adjacent Laz, which has prenominal relativization introduced by preverbal subordinators (Lacroix, this volume, chapter 6.2, §6.1).

- (49) (Gandon 2016: 220, citing Neocleous, in preparation, glosses modified)

- a. *i patsi [azi maziran d=erθen]*
 the.NOM girl.NOM from.the Mazira.ACC REL=came.3SG
d=emon d=anepsin eton
 the=my.NOM the=niece.NOM be.IPFV.3SG

²³ I interpret (47) as shown, but Gandon’s translation is slightly different.

- b. [asi maziran *d=erθen]* *i* *patsi*
 from.the Mazira.ACC REL=came.3SG the.NOM girl.NOM
d=emon *d=anepsin* *eton*
 the=my.NOM the=niece.NOM be.IPFV.3SG
 ‘The girl who came from Mazira is my niece’

In Istanbulite Romeyka, relativization may lack an overt relativizer if the head noun is present (50), and there appears to be a clitic resumptive pronoun on the relative clause verb (=na). In free relatives, a relativizer *op* is used (51–52) (see also Gandon 2016: 222 for free relatives), and with a spatial complement, the relativizer (*n*)*do* may be found (53). The verb in the relative clause appears to inflect for subject number, but we lack detailed information on the kinds of verb forms used here.

- (50) (Own fieldwork; ROF)

Aso Katoχor erθe o Mehmetis iða=na
 from.the Katochor came.3SG the Mehmet saw.1SG=3SG
 ‘I saw Mehmet who came from Katochor.’

- (51) (Own fieldwork; ROF)

Aso Katoχor op erθe iða=na
 from.the Katochor REL came.3SG saw.1SG=3SG
 ‘I saw the one that came from Katochor.’

- (52) (Own fieldwork; ROF)

So Istanbul op ine ul o Rumdža u=kserone.
 in.the Istanbul REL be.3PL all the Rumca NEG=know.3PL
 ‘Those (being) in Istanbul all don’t know Rumca.’

- (53) (Own fieldwork; ROF)

Sin džami etšika ndo en t=ospı temo en.
 at.the mosque down_there REL be.3SG the=house 1SG.POSS be.3SG
 ‘The house at the mosque is mine.’

6.5.1.2. Finite complement clauses in verbs of saying

Subordinate clauses introduced by verbs of saying have a finite verb form in the dependent clause which is either center-embedded as in relatives and embedded imperatives (54), (55), or post-posed (56), (57). Note that all the available examples could be interpreted as direct speech (cf. the second-person pronouns in the subordinate clause); to what extent direct and direct speech are differentiated syntactically remains an open question. With verbs of saying, in general no complementisers are used (Sitaridou 2014a: 128) but note the use of *ki* as complementiser in (57).

- (54) (Özkan, n.d.; RSür, glosses mine)

Eyo nt=eyapesa=se son kosmo ipa.
 I that=loved.1SG=you to.the.ACC world.ACC told.1SG
 ‘I have told the world that I loved you.’

- (55) (Sitaridou 2014a: 129; ROf, glosses modified)

O Alis esenan yrapson ipen.
 the.NOM Alis.NOM you write.IMP.2SG told.3SG
 ‘Alis ordered you to write.’

- (56) (Sitaridou 2014a: 128; ROf, glosses modified)

Tin patsi=m ipa espero=se yalemi.
 the.ACC daughter.ACC=my told.1SG take.1SG=you pencil.ACC
 ‘I told my daughter that I would buy her a pencil.’

- (57) (Own fieldwork; ROf)

Eyo ton džirim=m ipa ki bola eyabo=se.
 I the.ACC father.ACC=1SG told.1SG COMP much love.1SG=you
 ‘I told my father that I love him much.’

6.5.1.3. *na*-Clauses

This description of *na*-clauses is based on Sitaridou (2014a: 123–126). Apart from its function as a future marker, *na* is used as a subordinating conjunction (Mackridge 1987: 130). In comparison to SMG, its use in Romeyka is more restricted: *na* occurs in nonveridical predicates such as volitionals (58a) and negated present tense modals (58b), as well as in causatives (see ex. 45), mental perception verbs such as *enespala* ‘I forgot’ (58c) and *erte so tšefali=m* ‘it came to my mind’, and emotive verbs (58d) (Sitaridou 2014a).

- (58) (Examples from Sitaridou 2014a: 123–126; ROf, glosses modified)

- a. *Utš eθelna n' emaireva*
 not want.IPFV.1SG FUT cook.IPFV.1SG
 ‘I didn't want to cook.’
- b. *U poro n' almeyo.*
 not can.1SG FUT milk.1SG
 ‘I cannot milk (the cows).’
- c. *Enespala na leyo ti mami*
 forget.PST.1SG FUT say.1SG the grandmother.ACC
ta xaberae.
 the.ACC news.ACC
 ‘I forgot to tell the news to the grandmother.’

- d. *Exara na mairevo.*
 be_happy.PST.1SG FUT cook.1SG
 ‘I was happy I had cooked.’

Note in (58a) that volitional *θelo* only requires a *na*-clause as complement when negated. This does not hold true for the volitional *ayapo* ‘I love/like’ which always requires a *na*-clause (59) (Sitaridou 2014a: 124).

- (59) (Own fieldwork; ROF)
Eyo esena eyabo ne lepo.
 I you like.1SG FUT see.1SG
 ‘I want you to see (that).’

6.5.2. Non-finite subordination

6.5.2.1. Infinitives

According to Sitaridou (2014a: 122), the infinitive in Romeyka is defined as not bearing any agreement features, and as not occurring independently, but only as the complement of a superordinate syntactic construction (Özkan 2013: 149). Normally, the infinitive consists of the aorist stem followed by the infinitival ending *-ini* which can also bear the passive stem marker *-θ-*. The infinitive may have a distinct nominative subject such as in inflected and personal infinitives (see Sitaridou 2014b). Furthermore, it can take either a predicate or a DP as a complement, it can be coordinated, and it can be modified by adverbs (Sitaridou 2014a: 130).

Infinitives in Romeyka occur (i) in negated past tense modals (60a); (ii) as complements to the negated past tense volitional *utš eθelesa* ‘I didn’t want’ (60b); (iii) in *prin* ‘before’-clauses (60c) but not with other prepositions such as *os* ‘until’; (iv) in counterfactuals such as wishes, exclamatives, and conditionals as a complement of ‘have’ (60d) (Sitaridou 2014a, 2014b). Unlike in SMG, the infinitive in Romeyka does not occur in: (i) positive past tense modals, (ii) negated past tense modals, and (iii) modals in questions (Sitaridou 2014a: 119). Sitaridou (2014a) argues that the Romeyka infinitive behaves like a negative polarity item which is licensed by antiveridicality.

- (60) (Sitaridou 2014a: 126, 120; exs. 25, 26, 9a, 9b, respectively; ROF, glosses modified)

- a. *Utš eporesa tšimiθini.*
 not can.PST.1SG sleep.INF
 ‘I could not sleep.’
- b. *Utš eθelesa mairepsini.*
 not want.PST.1SG cook.INF
 ‘I didn’t want to cook.’

- c. *Prin spudzisini so mandarin, tsi pao.*
before clean.INF at.the barn not go.1SG
'I am not leaving before I clean the barn.'
- d. *As išen portpatesini sa rašia.*
OPT have.PST.3SG walk.INF to.the mountains
'S/He should have walked in the mountains.'

The syntactic constructions in which the infinitive occurs vary between dialects (Mackridge 1999: 102–103). For example, in RSür as spoken in Beşköy, the infinitive is not used in *prin* ‘before’-clauses (Özkan 2013: 149). Furthermore, in this variety, the infinitive is inflected by active past personal endings added to the infinitive when occurring after the negated past tense of *θelo* ‘I want’ and *poro* ‘I can’, like in *utš eθelesa porpatesna* ‘I did not want to walk’ (Özkan 2013: 148).

Finally, inflected infinitives occur as a strategy of nominalization whereby they are complements to (a) aspectuals such as *epiturepsa* ‘I finished’, and (b) verbs of mental perception such as *enespala* ‘I forgot’ (61). The nominalized infinitive is used with a, possibly obligatory, complex possessive (*e)muneθe* which may be a calque from Turkish nominalizations of the type *oku-ma-si-ni* ‘read-INF-3SG.POSS-ACC’ (Sitaridou 2014a: 130). Furthermore, nominalization by means of inflected infinitives before adjectives are formed as in (62, cf. also ex. 9) (Sitaridou 2014b: 42).

- (61) (Sitaridou 2014a: 131; ROF, glosses modified)
To tšimiθin=emun-eθe enespala.
 the sleep.INF=our-its forget.PST.1SG
 'I forgot to sleep.'
- (62) (Sitaridou 2014b: 42, Romeyka, glosses modified)
Afti i δulian to maθini=mu yola en
 this.NOM the.NOM job.NOM the learning=my easy be.3SG
 'It is easy for me to learn how to do this job.'

6.5.2.2. Deverbal nouns

Along with infinitives, Sitaridou (2014a) also recognizes verbal nouns. They may be used (a) as complements to volitionals such as *θelo* ‘I want’ (63a), and (b) as a complement to the phase modal *epašlaepsa* ‘I started’ (63b) which requires a deverbal noun introduced by the preposition *so* ‘to the’ (Sitaridou 2014a: 130).

- (63) (Sitaridou 2014a: 131, 130, respectively; ROF, glosses modified)
- a. *to peđi=m to panimon aso xorion θelo.*
the child=my the going from.the village want.1SG
'I want my child to leave the village.'
 - b. *Epašlaepsa pola so ðipsasimo.*
start.PST.1SG lot to.the drinking
'I started to get very thirsty.'

6.6. Clausal complements lacking an overt complementiser

The following is based on Sitaridou's (2014a: 127–129) discussion of null complementisers. Several types of complement clause occur without an overt complementiser. These include the complements of (i) perception verbs (64a); (ii) some emotive verbs such as *efoveθa* 'I feared' (64b) (but not in all emotive verbs; others are expressed by *na*-clauses); (iii) epistemic predicates (64c); (iv) verbs of saying (see ex. 56 in §6.5.1.2).

- (64) (Sitaridou 2014a: 127, 128; exs. 27, 30, 33 respectively; ROF, glosses modified)
- a. *Eyo ekusa o tšopanon ton arko endoke.*
I.NOM hear.PST.1SG the shepherd.NOM the.ACC wolf.ACC
kill.PST.3SG
'I heard that the shepherd killed the wolf.'
 - b. *Efoveθa xanis ton para=s.*
fear.PST.1SG lose the.ACC money.ACC=your
'I feared you may lose your money.'
 - c. *θaro xastas en.*
think.1SG sick be.3SG
'I think s/he is sick.'

6.7. Adverbs and adverbial clauses

6.7.1. Adverbial clauses of time and purpose

Sitaridou (2014a) discusses the use of the infinitive in *prin* 'before'-clauses (see ex. 60c in §6.5.2.1). Other temporal adjuncts, however, do not select an infinitive and require rather a *na*-clause (65). Temporal adverbial clauses are introduced by an adverbial subordinator which appears at the beginning of the subordinate clause, e. g. *anda* 'when' or *os* 'until' (65, 66).

- (65) (Sitaridou 2014a: 135; ROF, glosses modified)

Os na 'rte o Mehmetis na permeno=se.
until FUT come.3SG the.NOM Mehmet.NOM FUT wait.1SG=you
'I'll wait for you until Mehmet comes.'

- (66) (Sitaridou 2014b: 46; ROF, glosses modified)

Anda pašlaevis so fanimon os na piturevis ata,
when start.2SG at.the eating.ACC until FUT finish.2SG them.ACC.CL
u poris na stetšis.
not can.2SG FUT stop.2SG
'Once you start eating, you can't stop until you finish all of it.'

Adjuncts of purpose after verbs of motion show nominalization (67) (Sitaridou 2014b: 41).

- (67) (Sitaridou 2014b: 41; ROF, glosses modified)

ðæ te mon t' erθan=im, xavitsin efteſ me.
for my the coming=my pudding.ACC make.2SG me.ACC.CL
'For the sake of my coming, you will make me some pudding.'

6.7.2. Comparative Constructions

Comparative constructions in Romeyka have not been previously described. The standard of comparison seems to be introduced by the comparative particle *as* which has a locative function (68) (or *ap*, especially with anaphors (69), Drettas 1997). There is no case marking neither on the standard NP nor on the comparee NP.

- (68) (Özkan, n.d.; RSür, glosses mine)

o Mehmetis aso Osmanis alo tranos en.
the Mehmet from.the Osman more big be.3SG
'Mehmet is bigger than Osman.'

- (69) (Own fieldwork; ROF)

Temetero to χorio ap=esetero emorfi-tero en.
our the village PRT=your beautiful-CMP be.3SG
'Our village is more beautiful than yours.'

While (68) from RSür forms the comparison peripherastically with *alo* 'more', ROF uses the invariable(?) verbal comparative suffix *tero* (69) which is derived from AG verbal stem+*o*+*τερ*+inflectional ending. Furthermore, in Istanbulite Romeyka Turkish *fazla* 'more' is borrowed (70).

- (70) (Own fieldwork; ROF)

Talo fazla tar en.
 one more narrow be.3SG
 ‘One is narrower.’

The superlative is built peripherastically by *olon* ‘all’ (71). There seems to be variation with regard to as to whether the preposition *as* ‘from’ is used before *olon*.

- (71) (Own fieldwork; ROF)

Olon emorfo o=spi temo en.
 all beautiful the=house 1SG.POSS be.3SG
 ‘The nicest house is mine.’

6.8. Conditional clauses

Table 13 (based on Table 3 in Sitaridou 2014a: 122) summarizes some types of conditionals in ROF. See (72) for an example of a simple conditional clause.

Table 13: Conditionals in ROF (adapted from Table 3 in Sitaridou 2014a: 122)

		If-clause	Matrix clause
Nonveridical	Possible	<i>na</i> + present <i>an</i> + present	<i>na</i> + present (=future)
	Irrealis	<i>n'</i> + imperfect <i>an</i> + imperfect	<i>n'</i> + imperfect
Antiveridical (counterfactual)		<i>na/ŋ' + ixa</i> ‘I had’ + infinitive	
		<i>an + ixa</i> ‘I had’ + infinitive	<i>ixa</i> ‘I had’ + infinitive

The examples in (73) and (74) show some of the micro-variation in Romeyka conditionals. In RSür as spoken in Beşköy, the irrealis construction requires an inflected infinitive with an uninflected auxiliary. Sitaridou (2014a: 55) notes the possible emergence of a new invariant modality marker *ixe* from a former auxiliary.

- (72) (Sitaridou 2014a: 141; ROF)

An eporo, mairevo
 if can.1SG cook.1SG
 ‘If I can, I cook.’

- (73) (Sitaridou 2014b: 45; ROF, glosses modified)

An ixa mairepsini, n' etroyame
 if had.1SG cook.INF FUT eat.IPFV.1PL
 ‘If I had cooked, we would have eaten.’

- (74) (Sitaridou 2014b: 49; RSür, glosses modified, translation from Sitaridou)
- | | | | | |
|----------------------------------|---------|------------------|--------------|--|
| Na | ixa | mairepsina | etroyamen | |
| FUT | had.1SG | cook.AOR.INF.1SG | eat.IPFV.1PL | |
| ‘If I had cooked, we would eat.’ | | | | |

Abbreviations

ACC	accusative	N	neuter
AOR	aorist	NEG	negation
AUX	auxiliary	NMLZ	nominalization
CL	clitic	NOM	nominative
CMP	comparative	OPT	optative
COMP	complementiser	PASS	passive
DEF.ART	definite article	PL	plural
DEM	demonstrative	POSS	possessive
F	feminine	PRS	present
FOC	focalizer	PRT	particle
FUT	future	PST	past (this term subsumes e. g. “past perfect” of Sitaridou 2014a)
GEN	genitive		
HUM	human		
IMP	imperative	Q	question particle
INF	infinitive	REL	relativizer
IPFV	imperfective	SG	singular
M	masculine	1, 2, 3	1 st , 2 nd , 3 rd person

References

- Andrews, Peter A. 1989. *Ethnic groups in the Republic of Turkey*. Compiled and edited with the assistance of Rüdiger Benninghaus. Wiesbaden: Reichert.
- Bilici, Faruk. 2011. Que reste-t-il de la langue et de la culture grecques sur les côtes turques de la mer Noire? *Cahiers Balkaniques* 38–39. 325–342.
- Bortone, Pietro. 2009. Greek with no history, no standard, no models: Muslim Pontic Greek. In Alexandra Georgakopoulou & Michael Silk (eds.), *Standard Languages and Language Standards: Greek, Past and Present*, 67–89. London: Ashgate.
- Brendemoen, Bernt. 2002. *The Turkish dialects of Trabzon: Their phonology and historical development*. Wiesbaden: Harrassowitz.
- Brendemoen, Bernt. 2006. Aspects of Greek-Turkish language contact in Trabzon. In Hendrik Boeschoten & Lars Johanson (eds.), *Turkic Languages in Contact*, 63–73. Wiesbaden: Harrassowitz.
- Chatzopoulou, Katerina & Ioanna Sitaridou. 2014. Jespersen’s cycle for NEG2 and conditional inversion in the history of Greek: Evidence from Romeyka conditionals. Paper presented at the 16th Diachronic Generative Syntax Conference, Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, 3–5 July, 2014. http://www.nytud.hu/digs16/absz/chatzopoulou_sitaridou_abstract.pdf (accessed 13 November 2017).

- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Dawkins, Richard M. 1916. *Modern Greek in Asia Minor: A study of the dialects of Silli, Cappadocia and Pharasa with grammar, texts, translations and glossary*. Cambridge: Cambridge University Press.
- Dawkins, Richard M. 1931. Notes on the study of the Modern Greek of Pontos. *Byzantion* 6. 389–400.
- Dawkins, Richard M. 1937. The Pontic dialect of Modern Greek in Asia Minor and Russia. *Transactions of the Philological Society* 36. 15–52.
- Deffner, Michael. 1878. Die Infinitive in den pontischen Dialekten und die zusammengefügten Zeiten im Neugriechischen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, 191–230. Berlin: Buchdruckerei der Königlichen Akademie der Wissenschaften.
- Drettas, Georges. 1997. *Aspects pontiques*. Paris: Association de Recherches Pluridisciplinaires.
- Drettas, Georges. 1999. Το ελληνο-ποντιακό διαλεκτικό σύνολο [The Greek-Pontic dialect group]. In Anastasios-Foivos Christidis, Maria Arapopoulou & Giannoula Giannoulopoulou (eds.), *Διαλεκτικοί θύλακοι της ελληνικής γλώσσας* [Dialect enclaves of the Greek language], 91–100. Athens: Ministry of National Education and Religious Affairs/Centre for the Greek Language.
- Gandon, Ophélie. 2016. La relativisation dans une perspective aréale: l'aire Caucase – Anatolie de l'Est – Iran de l'Ouest. Paris: Université Sorbonne Nouvelle dissertation.
- Janse, Mark. 2004. Animacy, definiteness, and case in Cappadocian and other Asia Minor Greek dialects. *Journal of Greek Linguistics* 5. 3–26.
- Janse, Mark. 2008. Clitic doubling from Ancient to Asia Minor Greek. In Dalina Kallulli & Liliane Tasmowski (eds.), *Clitic Doubling in the Balkan Languages*, 165–202. Amsterdam & Philadelphia: Benjamins.
- Karatsareas, Petros. 2011. A study of Cappadocian Greek nominal morphology from a diachronic and dialectological perspective. Cambridge: University of Cambridge dissertation.
- Karatsareas, Petros. 2014. On the diachrony of gender in Asia Minor Greek: The development of semantic agreement in Pontic. *Language Sciences* 43. 77–101.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the World*, 19th edn. Dallas: SIL International. <http://www.ethnologue.com> (accessed 29 March 2016).
- Mackridge, Peter. 1987. Greek-Speaking Moslems of North-East Turkey: Prolegomena to study of the Ophitic sub-dialect of Pontic. *Byzantine and Modern Greek Studies* 11. 115–137.
- Mackridge, Peter. 1999. The Greek spoken in the Region of Pontus. In Anastasios-Foivos Christidis, Maria Arapopoulou & Giannoula Giannoulopoulou (eds.), *Διαλεκτικοί θύλακοι της ελληνικής γλώσσας* [Dialect enclaves of the Greek language], 101–105. Athens: Ministry of National Education and Religious Affairs/Centre for the Greek Language.
- Michelioudakis, Dimitris & Ioanna Sitaridou. 2012a. Multiple wh-fronting across Pontic Greek dialects. In Angela Ralli, Brian Joseph, Mark Janse & Athanasios Karasimos (eds.), *Proceedings of the International Conference of Modern Greek Dialects and Linguistic Theory* 5(1). 353–378.

- Michelioudakis, Dimitris & Ioanna Sitaridou. 2012b. Syntactic microvariation: Dative constructions in Greek. In Ricardo Etxepare & Beatriz Fernández (eds.), *Datives in variation: A micro-comparative perspective*, 212–255. Oxford: Oxford University Press.
- Moseley, Christopher. 2007. *Encyclopedia of the World's Endangered Languages*. London: Routledge.
- Moseley, Christopher. 2010. *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing. <http://www.unesco.org/new/en/communication-and-information-access-to-knowledge/linguistic-diversity-and-multilingualism-on-internet/atlas-of-languages-in-danger/> (accessed 17 November 2017).
- Neocleous, Nicolaos. 2017. The reconstruction of VO and OV alternation in Romeyka. Paper presented at the Conference on Language Contact and Language Change in Western Asia, Goethe University Frankfurt, 10–12 March, 2017.
- Neocleous, Nicolaos. In preparation. Headedness in Romeyka: Continuity, contact and change in the Hellenic varieties of Pontus. Cambridge: University of Cambridge dissertation.
- Oikonomidis, Demosthenes E. 1908. *Lautlehre des Pontischen*. Leipzig: A. Deichert'sche Verlagsbuchhandlung Nachf.
- Özkan, Hakan. 2013. The Pontic Greek spoken by Muslims in the villages of Beşköy in the province of present-day Trabzon. *Byzantine and Modern Greek Studies* 37(1). 130–150.
- Özkan, Hakan, n.d. Grammar sketch of Romeyka. Unpublished draft.
- Parcharidis, Ioannis. 1880. *Γραμματική της διαλέκτου Τραπεζούντος: Περί επιθέτων, αντωνυμιών και βαρυτόνων ρημάτων*. [Grammar of the dialect of Trabzon]. Χειρόγραφο 335, Κέντρο Ερεύνης των Νεοελληνικών Διαλέκτων και Ιδιωμάτων, Ακαδημία Αθηνών.
- Parcharidis, Ioannis. 1888. *Συλλογή ζόντων μνημείων της αρχαίας ελληνικής γλώσσας εν Οφει*. [Collection of living documents of the Ancient Greek language in Ophis] *O εν Κωνσταντινούπολει Ελληνικός Φιλολογικός Σύλλογος* 18 (1883–1884). 120–178.
- “Pontic”. *Catalogue of Endangered Languages*. 2017. University of Hawaii at Manoa. <http://www.endangeredlanguages.com/lang/3265> (accessed 17 November 2017).
- Schreiber, Laurentia. 2016. Assessing sociolinguistic vitality: An attitudinal study of Rumca (Romeyka). In Heike Wiese, Heiko F. Marten & Oliver Bunk (eds.), *Arbeitspapiere "Sprache, Variation und Migration": Studentische Arbeiten, Papier Nr. 4*. Potsdam: Universität Potsdam. https://www.unipotsdam.de/fileadmin01/projects/svm/Arbeitspapiere/MA_Schreiber_May.pdf (accessed 17 November 2017).
- Schreiber, Laurentia & Ioanna Sitaridou. 2017. Assessing the sociolinguistic vitality of Istanbulite Romeyka: An attitudinal study. *Journal of Multilingual and Multicultural Development* 39(1). 1–16. <http://www.tandfonline.com/doi/abs/10.1080/01434632.2017.1301944?journalCode=rmmm20> (accessed 17 November 2017).
- Sitaridou, Ioanna. 2013. Greek-speaking enclaves in Pontus today: The documentation and revitalization of Romeyka. In Mari Jones & Sarah Ogilvie (eds.), *Keeping languages alive: Documentation, pedagogy, and revitalization*, 98–112. Cambridge: Cambridge University Press.
- Sitaridou, Ioanna. 2014a. Modality, antiveridicality, and complementation: The Romeyka infinitive as a negative polarity item. *Lingua* 148. 118–146.
- Sitaridou, Ioanna. 2014b. The Romeyka infinitive: Continuity, contact and change in the Hellenic varieties of Pontus. *Diachronica* 31(1). 23–73.

7. Appendix: Basic lexical items in selected languages of Western Asia

Geoffrey Haig, Geoffrey Khan and Laurentia Schreiber

7.1. Background

The list of items is based on a list of 40 meanings (a sub-set of the Swadesh-100 list), that is used by the *Automated Similarity Judgement Program* (ASJP), developed by Søren Wichmann and associates (Wichmann et al. 2016). The program is designed for automatically classifying languages according to their origins, and for inferring dates of divergences of different groups (see Holman et al. 2011, and accompanying peer commentaries for critical assessment of the method). The compilation includes lists from several chapters in this volume (provided by the respective authors), as well as lists provided from other sources. In §7.5 below, we provide a list of sources with accompanying notes on transcriptions etc.

For some meanings, no forms were provided due to e. g. lack of a reliable response for the item concerned. The relevant cells in the Tables have been left empty. The list of meanings is arranged alphabetically, except for the personal pronouns, which are grouped together at the end of the table. Following a suggestion by Don Stilo, an additional four items were added to the Wichmann-list ('cat', 'water-melon', 'frog', 'why?') which are known to provide interesting lexical isoglosses in the languages of the region.

Please note that this is primarily a compilation of lexical items, for use in e. g. recognizing cognates, or borrowings. It does not claim phonetic accuracy or consistency in the way these items are rendered in the different lists. Given the very heterogenous nature of the sources, and the differing traditions in the different disciplines represented in this compilation, imposing a standardized transcription would have not have been practicable, and we have therefore left the transcriptions in the manner they were provided to us. For further information on the sources and conventions for the individual lists, please refer to the accompanying notes under §7.5 below.

7.2. Iranian languages

	Behd.	Gor.I	Gor.II	Hawr.	S. Kurd.	Bakht.I	Bakht.II
BLOOD	<i>xīn</i>	<i>xīn</i>	<i>χʷiin</i>	<i>wiñi</i>	<i>xʷe:n</i>	<i>hin</i>	<i>xin</i>
BONE	<i>hesī</i>	<i>pēša</i>	<i>soqan</i>	<i>pēša</i>	<i>soqan</i>	<i>hast, ostoxovn</i>	<i>estexun</i>
BREAST	<i>sīng</i>		<i>sinæ</i>	<i>sīna</i>	<i>sinæ</i>	<i>pestovn, mame</i>	<i>pestun, mame</i>
CAT	<i>pišk</i>		<i>pʃi</i>	<i>kite</i>	<i>pʃi</i>	<i>gorbe</i>	<i>gorbe</i>
COME	<i>dihēt</i>	<i>mayo</i>	<i>miyo</i>	<i>āmāy</i>	<i>eyye:, æte:</i>	<i>oveyden</i>	<i>oveydan</i>
DIE	<i>dimirīt</i>	<i>mamro</i>	<i>min⁹ru</i>	<i>marday</i>	<i>emre:</i>	<i>morden</i>	<i>mordan</i>
DOG	<i>se</i>	<i>sipa</i>	<i>sipæ, gæmal</i>	<i>tūta</i>	<i>sew</i>	<i>say</i>	<i>sag</i>
DRINK	<i>vedixot</i>	<i>moro</i>	<i>mt̪itru</i>	<i>wārday</i>	<i>æxwa</i>	<i>xarden</i>	<i>xardan</i>
EAR	<i>guh</i>		<i>guʃ</i>	<i>goš</i>	<i>gijikæ</i>	<i>guš</i>	<i>guš</i>
EYE	<i>čav</i>	<i>čam</i>	<i>ʃem</i>	<i>čam</i>	<i>ʃaw</i>	<i>tey</i>	<i>tiye</i>
FIRE	<i>agir</i>	<i>ārir</i>	<i>awir</i>	<i>ayr</i>	<i>awir</i>	<i>taš</i>	<i>taš</i>
FISH	<i>masī</i>	<i>māšāw</i>	<i>masew</i>	<i>māšāw</i>	<i>mahi</i>	<i>māi</i>	<i>māi</i>
FROG	<i>beq</i>		<i>qurwakə</i>	<i>qurwāqə</i>	<i>qaq</i>	<i>baq</i>	<i>baq</i>
FULL	<i>tījī</i>		<i>pir</i>	<i>pař</i>	<i>pir</i>	<i>por</i>	<i>por</i>
HAND	<i>dest</i>	<i>dast</i>	<i>des</i>	<i>das</i>	<i>des</i>	<i>dast</i>	<i>dast</i>
HEAR	<i>gulē dibīt</i>		<i>miʒnāw</i>	<i>ēga</i>	<i>æznaeʃe:</i>	<i>ašniden</i>	<i>fahmidan</i>
HORN	<i>šax</i>		<i>fax</i>	<i>šax</i>	<i>fax</i>	<i>šāx</i>	<i>šāx</i>
KNEE	<i>čok</i>	<i>zirani</i>		<i>zangol</i>	<i>wižin</i>	<i>zuni</i>	<i>zuni</i>
LEAF	<i>belk</i>	<i>walig</i>	<i>bærg, glaw</i>	<i>galā, walg</i>	<i>gæta</i>	<i>par</i>	<i>pahr</i>
LIVER	<i>mēlak</i>		<i>đier</i>	<i>yahar; jarg</i>	<i>đie;r</i>	<i>jivar</i>	<i>jivar</i>

	Behd.	Gor.I	Gor.II	Hawr.	S. Kurd.	Bakht.I	Bakht.II
LOUSE	<i>sپیھ</i>			<i>hašpšē</i>	<i>sipe:</i>	<i>šeš</i>	<i>šeš</i>
MOUNTAIN	<i>cǐya</i>	<i>kūh</i>	<i>ku(h)</i>	<i>kaš-u-ko</i>	<i>ke:f</i>	<i>koh</i>	<i>koh</i>
NAME	<i>nav</i>	<i>nāmo</i>	<i>nam</i>	<i>namē</i>	<i>naw</i>	<i>esm, nom</i>	<i>esm, nom</i>
NEW	<i>nī</i>		<i>tazē</i>	<i>tāza</i>	<i>tażə</i>	<i>nu</i>	<i>nav</i>
NIGHT	<i>šev</i>	<i>šaw</i>	<i>ʃaw</i>	<i>šawē</i>	<i>ʃəw</i>	<i>šav</i>	<i>šav</i>
NOSE	<i>dīfin</i>	<i>pit</i>	<i>pit</i>	<i>lūta</i>	<i>lūt</i>	<i>noft</i>	<i>neft</i>
ONE	<i>ēk</i>	<i>īka</i>	<i>yək</i>	<i>yak</i>	<i>yək</i>	<i>yak</i>	<i>yek, yak</i>
PATH	<i>rēk</i>	<i>rēga</i>	<i>ra</i>	<i>rā</i>	<i>re:</i>	<i>rah, jahde</i>	<i>rah, jāde</i>
PERSON	<i>kes</i>		<i>kes</i>	<i>kas</i>	<i>keṣ</i>	<i>ādom</i>	<i>ādam</i>
SEE	<i>dibinū</i>	<i>matū</i>	<i>mi:nu</i>	<i>dīyay</i>	<i>čeünē:</i>	<i>dīden</i>	<i>dīdan</i>
SKIN	<i>cerm, cild</i>	<i>posta</i>	<i>pus</i>	<i>pos</i>	<i>pus</i>	<i>pust</i>	<i>pust</i>
STAR	<i>stēr</i>	<i>hasṭāra</i>	<i>asarae</i>	<i>hasāra</i>	<i>hesa:ræ</i>	<i>āstāre</i>	<i>āstāre</i>
STONE	<i>ber</i>	<i>kamar</i>	<i>kəmær</i>	<i>twanī</i>	<i>kofik</i>	<i>bard, kočok</i>	<i>bard</i>
SUN	<i>roj</i>	<i>hor</i>	<i>xwær</i>	<i>waratāw, war,</i>	<i>xweer</i>	<i>aftav</i>	<i>aftav</i>
TONGUE	<i>ezman</i>	<i>zwāñ</i>	<i>zʷwan</i>	<i>zwāñ</i>	<i>zwan</i>	<i>zovn</i>	<i>zun</i>
TOOTH	<i>didan</i>	<i>diđān</i>	<i>diyan</i>	<i>diđān</i>	<i>diyan</i>	<i>dendom</i>	<i>dendum</i>
TREE	<i>dar</i>	<i>dār</i>	<i>dar</i>	<i>dār, diraxt</i>	<i>dar</i>	<i>dār</i>	<i>deraxt</i>
TWO	<i>dū</i>	<i>duwa</i>	<i>du</i>	<i>dwē</i>	<i>do</i>	<i>do</i>	<i>do</i>
WATER	<i>av</i>	<i>ā</i>	<i>aw</i>	<i>āwī</i>	<i>aw</i>	<i>av</i>	<i>av</i>

	Behd.	Gor.I	Gor.II	Hawr.	S. Kurd.	Bakht.I	Bakht.II
WATER-MELON	šiū	<i>fani</i>	<i>hamī</i>	<i>fāni</i>	<i>hendowne</i>	<i>hendune</i>	
WHY?	bočī	<i>bila</i>	<i>payčī</i>	<i>buʃʃe, æra</i>	<i>siče</i>	<i>siče</i>	
I	ež, min	<i>amin</i>	<i>amin</i>	<i>min</i>	<i>mo</i>	<i>mo</i>	
WE	em, me	<i>ēma</i>	<i>imæ</i>	<i>imæ</i>	<i>imā</i>	<i>imā</i>	
YOU (SG)	tu, te	<i>ati</i>	<i>to</i>	<i>to</i>	<i>to</i>	<i>to</i>	
YOU (PL)	hīn, hīngo	<i>šima</i>	<i>ſimæ</i>	<i>ſima</i>	<i>ivæ</i>	<i>isā</i>	<i>šomā</i>

	Kumz.¹	Oss.	Casp.I	Casp.II	Tal.I	Tal.II	Tat.I	Tat.II
BLOOD	xwəm	<i>tug/tog</i>	<i>xun</i>	<i>xun</i>	<i>xun</i>	<i>xun</i>	<i>xun</i>	<i>xun</i>
BONE	xār	<i>šneg/nsteg</i>	<i>xaš</i>	<i>hesseka</i>	<i>astæ</i>	<i>xaš</i>	<i>xeš</i>	<i>axdoxan</i>
BREAST	sīna	<i>riw/rew</i>	<i>čučay</i>	<i>juju</i>	<i>p̥išton*</i>	<i>jluju[#]</i>		<i>čize</i>
CAT	gurbač	<i>gədətikis</i>	<i>pič</i>	<i>bameši</i>	<i>pišik</i>	<i>puču</i>	<i>peču</i>	<i>kiğae</i>
COME	āmad/āmas/byāt/tī	<i>arba-səw-/vrba-cəw-</i>	<i>a/ama</i>	<i>e/amu</i>	<i>(v)o/o'me</i>	<i>a/om</i>	<i>a/ami</i>	<i>a-/ave</i>
DIE	murd/murs/mur/tumr	<i>mel-</i>	<i>merd</i>	<i>merd</i>	<i>merd</i>	<i>mer/mærd</i>	<i>mer/d</i>	<i>mer/mærd</i>
DOG	sag	<i>kʷəz/kuj</i>	<i>səg</i>	<i>sek</i>	<i>sipæ</i>	<i>espæ</i>	<i>espæ</i>	<i>espæ</i>
DRNK	šaraba gild/š. gis/s. kē/s. tka	<i>nwaž-/nivaz-</i>	<i>xor/xord</i>	<i>xard</i>	<i>pe-šom</i>	<i>xənj/xənt</i>	<i>henj/</i> <i>hent</i>	<i>der-ašow</i>
EAR	gōš	<i>quš/kos</i>	<i>guš</i>	<i>guš</i>	<i>guš</i>	<i>guš</i>	<i>guš</i>	<i>guš</i>

¹ See §7.5 below for selected Kumzari etymologies.

	Kumz. ¹	Oss.	Casp.I	Casp.II	Tal.I	Tal.II	Tat.I	Tat.II
EYE	čum	səšt/cəst	čušm	češ	čaš	čem	čæšm	čæmm
FIRE	ātiš	art	aťaš	taš	oťaš	ateš	ateš	ayr
FISH	møy	kešag, keʃʃəsalge	møy	møy	moy	moy	moy	mahi
FRG	šyfrāqō	χyʃʃ'χyʃʃə	yurhayə	vek	vəzəx	guzga	guzga	qurbaye
FULL	palla	zag/itəg	pər	per	pur	pər	fet	perr
HAND	dist	k'uχ, 'arm'/k'oχ	dəs	dəs	dəs	dəs	dəs	dəs
HEAR	šnuʃ/šnawd/šnavs/ šnēw	quš-/ibos-	ištowšən	šnuss	maes/e	maes/ə	mass	æšnæv
HORN	qarn	šək'a/šík'a	šax	šax	šox	šaγ [#]	šax	šax
KNEE	rulkbit	wrag/zangv	zonə	zanu	zonj	zongə	zeng	ərzane
LEAF	warq	šəʃʃifə	bəlg, vəlg	vəlg	livə	lev	liv	vælg
LIVER	jōğur	jiger	jigər	jiger	jigær	jigær	jiger	jiger
LOUSE	qar'a	šətʃ'sistə	subuj	espic	sihiž	əspəj	esbej	esbəzə
MOUNTAIN	aql	χɔχ'χ'wənχ	ku	ku	band	ku	ku	ku
NAME	nām	nom	nom	es	nom	nom, esm	nam	nawə
NEW	nō	nog/nəwəg	tazə	nu, taze	tožə	tazə	tazə	nu
NIGHT	šaw	əʃəv/əʃəvəv	šow	šu	šæv	šæv	šæv	šo
NOSE	nuxrit	fənz/fiʃ	dəmay	feni	vəni	vini	veni	viniæ
ONE	tā	ıw/jew	yek, i	yek	i	i	i	yey
PATH	tēra	fendag	ra	ru	ro	ra	ra	ra

	Kumz. ¹	Oss.	Casp.I	Casp.II	Tat.I	Tat.II	Tat.I	Tat.II
PERSON	ādimī	adžmag	adəm	adəm	adəm	adəm	adəm	adəm
SEE	jīr/jīš/māš/tēniš	fēn-/vīn-	dīn/dī	vīn/vīnd	vēn/vēnn	vīn/vīnd	vīn/dī	vīn/dī
SKIN	jīld, pōṣṭ	c'ar/c'arə	pūst	pūs	pōs	pūs	pūs	pūsd
STAR	starg	(ə)št'að/(ə)st'alu	sətarə	astare	setare	assarə	esdāre	
STONE	bard	dūr/dor	səng	səng	sīy	səng [#]	səg	səng
SUN	intaf	xur/xor	xoršīd	aftab	heši	xoršīd	aftav	xor
TONGUE	żwāñ	vvžag/vvzag	żebon	żebun	żivon	zebn	zuan	zuan
TOOTH	dnāñ	dəndag	dəndan,	dəndun	dandon	gaz	gaz	dəndan
			gaz					
TREE	šidrit	belaš/beləsə	dar	dar	do	dar	dar	dar
TWO	dō	dəwwəw/dəwwwə	du	do	di	də	de	do
WATER	āw	don	aw, ab	u	ov	av	av	āwəw
WATER- MELON	gahh	xarhəž/xarbuž (Turkish)	hindonə	hendev-	zimistoni	hindunə [#]	xendunə	henduane
WHY?	pi či	semən/cəmen (see notes)	čärə	če	boči(ro)	əčärə	čéra	ča
I	mē	vž/vz (NOM), mən (OBL)	mu	men	az, mi(n)	əz, mən	əz, mən	əz
WE	mā	maχ (NOM, OBL)	amo	ama, emə	əməc	əməc	əməc	awən
YOU (SG)	tō	də/dū (NOM), dəw (OBL)	tə	te	ti	tə	te	te
YOU (PL)	šmā	šəməχ/suməχ (NOM, OBL)	šəmo,	šəma	šiməc	šeməc	šeməc	soan

7.3. Semitic languages

	Arab.I	Arab.II	Arab.III	Arab.IV	CNA	NENA.I	NENA.II	NENA.III	NM
BLOOD	<i>damm</i>	<i>damm</i>	<i>damm</i>	<i>dam</i>	<i>'admo</i>	<i>dəmma</i>	<i>dəmma</i>	<i>dəma</i>	<i>dəmā</i>
BONE	<i>ṣaḍam</i>	<i>ṣaḍim</i>	<i>ṣaḍəm</i>	<i>ṣadum</i>	<i>garmo</i>	<i>farma</i>	<i>germa</i>	<i>garmā</i>	<i>gernā</i>
BREAST	<i>sədər</i>	<i>ṣadir;</i> female	<i>ṣadḡ;</i> female	<i>sadur</i>	<i>ṣadro</i>	⁺ <i>sadra;</i> female	<i>ṣadra;</i> female	<i>səŋga</i> (Kurd.); female	<i>kappa;</i> female breast: <i>nannā</i>
		<i>tadi</i>	<i>breast:</i>			<i>bəzzə</i>	<i>čəčča</i>		
CAT	<i>sannōra</i>	<i>bisse</i>	<i>bazzūni</i>	<i>bazzūna</i>	<i>qāṭen</i>	⁺ <i>kaṭu</i>	<i>qaṭu</i>	<i>qaṭu</i>	<i>šəṭārā</i>
COME	PFV <i>ḡā</i> (IPFV)	<i>ḡī</i> (PFV)	<i>ḡā</i> (PFV)	<i>ṛiyya</i>	<i>'oθe</i>	⁺ <i>atə</i>	⁺ <i>aθe</i> (PRS)	<i>he</i> (PRS)	<i>āṭi</i> (PRS)
	<i>yəḡī</i>	<i>yəḡī</i> (IPFV)	<i>yəḡī</i> (IPFV)	(PFV)	(PRS)				
DIE	<i>māt</i> (PFV)/ <i>ymūt</i> (IPFV)	<i>māt</i> (PFV)/ <i>ymūt</i> (IPFV)	<i>māt</i> (PFV)/ <i>ymūt</i> (IPFV)	(PFV)	<i>mayaθ</i>	<i>mayaθ</i>	<i>mayaθ</i> (PRS)	<i>mel</i> (PRS)	<i>møyet</i> (PRS)
DOG	<i>kalb</i>	<i>čalib</i>	<i>kalb</i>	<i>čaləb</i>	<i>kalbo</i>	<i>calba</i>	<i>kalba</i>	<i>kalba</i>	<i>kalbā</i>
DRINK	<i>ṣərəb</i> (PFV)/ <i>yəṣrəb</i> (IPFV)	<i>ṣrib</i> (PFV)/ <i>yiṣrab</i> (IPFV)	<i>ṣəgəb</i> (PFV)/ <i>yəṣğab</i> (IPFV)	(PFV)	<i>ʃote</i>	<i>ṣatə</i>	<i>ṣate</i> (PRS)	<i>ṣate</i> (PRS)	<i>ṣaṭjet</i> (PRS)
					(PRS)				
EAR	<i>ṛəḍən</i>	<i>ṛiḍin</i>	<i>ḍān</i>	<i>ḍān</i>	<i>'aðno</i>	<i>nata</i>	<i>naθa</i>	<i>nāḥalta</i>	<i>omnā</i>
EYE	<i>ṣayn</i>	<i>ṣēn</i>	<i>ṣēn</i>	<i>ṣēn</i>	<i>'ayno</i>	⁺ <i>'ayna</i>	<i>'ena</i>	<i>'ena</i>	<i>ina</i>
FIRE	<i>nār</i>	<i>nār</i>	<i>nāg</i>	<i>nār</i>	<i>nuro</i>	<i>muya</i>	<i>nura</i>	<i>nura</i>	<i>nurā</i>
FISH	<i>samake</i>	<i>ṣabbūt</i>	<i>samaki</i>	<i>samča</i>	<i>nənto</i>	<i>muya</i>	<i>nuna</i>	<i>nunila,</i> <i>nunilta</i>	<i>foniđā</i>

	Arab.I	Arab.II	Arab.III	Arab.IV	CNA	NENA.I	NENA.II	NENA.III	NM
FROG	<i>qaqroqa</i>	<i>quguruya</i>	<i>qaqoqa</i> (<* <i>qaqgōqa</i>)	<i>agruqa</i>	<i>bagqe</i>	<i>pəkkā</i>	<i>pəqqā,</i> <i>pəqqē</i>	<i>qurbaga</i> (Pers.)	<i>paqetā</i>
FULL	<i>məlli</i>	<i>malyān</i>	<i>məbabā</i>	<i>matriūs</i>	<i>məlyā</i>	<i>məlyā</i>	<i>zmaṭā</i>	<i>məli</i>	
HAND	<i>ṛīd</i>	<i>ṛīd</i>	<i>ṛīd</i>	<i>ṛīd</i>	<i>iḍo</i>	<i>iḍa</i>	<i>īla</i>	<i>īla</i>	<i>iḍā</i>
HEAR	<i>səməf</i> (Pfv)/ <i>yəsməf</i> (IPfv)	<i>smif</i>	<i>səməf</i> (Pfv)/ <i>yəsməf</i> (IPfv)	<i>simaſ</i> (Pfv)	<i>šomə</i> (PRS)	<i>šammo</i> (PRS)	<i>šāme</i> (PRS)	<i>šāvet</i> (PRS)	
HORN	<i>qərn</i>	<i>garin</i>	<i>qəğən</i>	<i>girin</i>	<i>qarno</i>	<i>kana</i>	<i>šaxa</i>	<i>qarn</i>	
KNEE	<i>rəgbə</i>	<i>rukba</i>	<i>rəkbī</i>	<i>rəkba</i>	<i>barko</i>	<i>bərca</i>	<i>bərka</i>	<i>bərka</i>	<i>burkā</i>
LEAF	<i>waraqa</i>	<i>wruqa</i>	<i>waqqa</i>	<i>wriga</i>	<i>tarfo</i>	<i>tarpa</i>	<i>tarpa</i>	<i>gala</i>	<i>par</i> (Pers.)
LIVER	<i>qasābe</i> (for dish also the Turkish loan <i>ğīgar</i>)	<i>čabde</i>	<i>kəbed</i>	<i>čabda</i>	<i>käzäbe</i>	<i>jifar</i>	<i>kawda</i>	<i>koza</i>	<i>kabdā</i>
LOUSE	<i>qamla</i>	<i>gamlı</i>	<i>qamli</i>	<i>gamla</i>	<i>qalmo</i>	<i>kalma</i>	<i>qalma</i>	<i>qalma</i>	<i>qamla</i>
MOUNTAIN	<i>ğabal</i>	<i>ğbile</i>	<i>ğabal</i>	<i>yibal</i>	<i>turo</i>	<i>ṭuṣra</i>	<i>tura</i>	<i>tura</i>	<i>tura</i>
NAME	<i>ṛəsəm</i>	<i>ṛəsəm</i>	<i>ṛisim</i>	<i>ṛisim</i>	<i>ṛəsmo</i>	<i>ṛəmma</i>	<i>ṛəmma</i>	<i>ṛəma</i>	<i>eṣmā</i>
NEW	<i>ğidād</i>	<i>ğidād</i>	<i>ğidād</i>	<i>yidād</i>	<i>ḥaθo</i>	<i>xata</i>	<i>xata</i>	<i>xala</i>	<i>haṭā</i>
NIGHT	<i>layl</i>	<i>lēl</i>	<i>lēl</i>	<i>lələ</i>	<i>laho</i>	<i>lele</i>	<i>lele</i>	<i>lilyā</i>	

	Arab.I	Arab.II	Arab.III	Arab.IV	CNA	NENA.I	NENA.II	NENA.III	NM
NOSE	<i>ʔanf</i>	<i>xušše</i>	<i>xašəm</i>	<i>darub</i>	<i>nħire</i>	<i>naxira</i>	<i>poqa</i>	<i>nħirā</i>	
ONE	<i>nħħad</i>	<i>nħħad</i>	<i>nħħad</i>	<i>nħħid</i>	<i>ħa (M), ħħo (F)</i>	<i>xa</i>	<i>xa</i>	<i>ħħa</i>	
PATH	<i>dərb</i>	<i>dariħ</i>	<i>daġħb (~tarīq)</i>	<i>darob</i>	<i>'urxa</i>	<i>'urxa</i>	<i>'orxa</i>	<i>ohrā</i>	
PERSON	<i>šaqṣ</i>	<i>šaxaṣ</i>	<i>šaxeṣ</i>	<i>nafar</i>	<i>zläm</i>	<i>naša</i>	<i>naša</i>	<i>bañnäšä</i>	
SEE	<i>ħatra (Pfv)/ <i>yəra</i> (IPfv)</i>	<i>ħäf</i> (Pfv)/ <i>yħuff</i> (IPfv)	<i>qəħəs'</i> (Pfv)/ <i>yəqħaṣ</i> (IPfv)	<i>ħäf</i> (Pfv)	<i>ħoże</i> (PRS)	<i>xazə</i> (PRS)	<i>xaze</i> (PRS)	<i>xäe</i> (PRS)	<i>ħäzi</i> (PRS)
SKIN	<i>għald</i>	<i>għilid</i>	<i>għelad</i>	<i>yillid</i>	<i>galdo</i>	<i>għolda</i>	<i>mäška</i>	<i>meškā</i>	
STAR	<i>näġme</i>	<i>naġme</i>	<i>naġni</i>	<i>nayma</i>	<i>kækwo</i>	<i>cuxxa</i>	<i>kaxxa</i>	<i>kxewla</i>	<i>kollha</i>
STONE	<i>ħaġar</i>	<i>ħaġra</i>	<i>ħaġag (~ħiġaġa)</i>	<i>saxra</i>	<i>kefø</i>	<i>cipa</i>	<i>kepa</i>	<i>għalta</i>	
SUN	<i>ħams</i>	<i>ħams</i>	<i>ħaməs</i>	<i>ħaməs</i>	<i>ħemra</i>	<i>ħemra</i>	<i>ħemra</i>	<i>ħemra</i>	<i>ħameš</i>
TONGUE	<i>lsēn</i>	<i>lsān</i>	<i>lsēn</i>	<i>ṛiħsān</i>	<i>lišono</i>	<i>lišana</i>	<i>lišana</i>	<i>lišana</i>	<i>lešnā</i>
TOOTH	<i>qərs</i>	<i>sinn</i>	<i>sənn</i> (~snēn)	<i>ðərəs</i>	<i>'aršo</i>	<i>cica</i>	<i>kaka</i>	<i>kaka</i>	<i>šenā</i>
TREE	<i>səġara</i>	<i>səġara</i>	<i>səġġa</i>	<i>fijra</i>	<i>dawmo</i>	<i>'ilana</i>	<i>'ilana</i>	<i>'ilana</i>	<i>dərxt</i> (Pers.)
TWO	<i>tnaqn</i>	<i>riħnene</i>	<i>imnēn</i>	<i>θniex</i>	<i>tre</i>	<i>tre</i>	<i>tre</i>	<i>tre</i>	<i>tren</i>
WATER	<i>mawj</i>	<i>mawj</i>	<i>māy</i>	<i>mawje</i>	<i>miyya</i>	<i>miyya</i>	<i>mae</i>	<i>mienn</i>	

	Arab.I	Arab.II	Arab.III	Arab.IV	CNA	NENA.I	NENA.II	NENA.III	NM
WATER - MELON	ḡəbse	dibše	šəmzi	rəjjiya	žabaš	šəfiyya (Kurd.)	šəfiyya (Kurd.)	šwətya (Kurd.)	qarambā
WHY?	layš	ṣalēš	lēš	li̥ʃ	qaŷ	ka-mu	qa-mo	ta-ma	qamu
I	zana	zāni	zana	zāna	'ono	'ana	'ana	'ana	anā
WE	nəħne	żihna	nəħna	żihna	'ahna	'axnan	'axni	'axni	ani
YOU (SG)	żənt (M) żənti (F)	żinti (M) żinti (F)	żənta (M) żənti (F)	żinta (M) żinti (F)	hat	'at, 'atən	'ati	'āt	āt
YOU (PL)	żəntən	żintu (M) żintin (F)	żəntən	żintum (M)	ħatu	'axtun, axtoxun, 'axnoxun	'axtu	'axtu	atton

7.4. Other language families

	Tk. Qorveh	Laz	Rom.
BLOOD	<i>qan</i>	<i>ditsxiri</i>	(<i>to</i>) (<i>j</i>) <i>øma</i>
BONE	<i>sümk</i>	<i>ili</i>	<i>ostuð</i>
BREAST	<i>döf</i>	<i>gjonksi</i>	(<i>ta</i>) <i>gønksa</i> (PL)
CAT	<i>pifik</i>	<i>k'at'u</i>	(<i>i</i>) <i>kata</i>
COME	<i>jelir</i>	<i>mo-xt-im-u</i>	<i>erχume</i>
DIE	<i>ölir</i>	<i>o-yur-u</i>	<i>χame</i>
DOG	<i>köpæk</i>	<i>dʒoyori</i>	<i>škilo</i>
DRINK	<i>itfir</i>	<i>o-f-u</i>	<i>pino</i>
EAR	<i>qulax</i>	<i>udʒi</i>	(<i>to</i>) <i>otiam</i> (possibly POSS.1SG)
EYE	<i>göz</i>	<i>toli</i>	(<i>ta</i>) <i>mata</i> (PL)
FIRE	<i>?ɔ:t</i>	<i>datʃxuri</i>	<i>apsomo</i>
FISH	<i>mahi</i>	<i>nɪʃxomi</i>	<i>opsar</i>
FROG	<i>qurbaæ</i>	<i>mai</i>	<i>furno</i>
FULL	<i>dolæ</i>	<i>opfa</i>	<i>eromoθe</i>
HAND	<i>?ael</i>	<i>xe</i>	(<i>to</i>) <i>šeri</i>
HEAR	<i>?eʃædir</i>	<i>o-gn-u</i>	<i>akuyo</i>
HORN	<i>ʃax</i>	<i>nkra</i>	<i>tšerato</i>
KNEE	<i>diz</i>	<i>burguli</i>	(<i>to</i>) <i>yonaton</i>
LEAF	<i>yarpax</i>	<i>butk'a</i>	<i>filo</i>
LIVER	<i>dʒægær</i>	<i>utʃa-f gui</i> (lit. ‘black heart’)	
LOUSE	<i>bit</i>	<i>mt'i</i>	<i>ftira</i>
MOUNTAIN	<i>dax</i>	<i>germa</i>	<i>raši(n)</i>
NAME	<i>?ad</i>	<i>dʒoxo</i>	(<i>to</i>) <i>onema</i>
NEW	<i>tazæ</i>	<i>ayne</i>	<i>jengi</i> (Turkish)
NIGHT	<i>gedʒæ</i>	<i>seri</i>	<i>opse</i>
NOSE	<i>birn</i>	<i>tʃindi</i>	(<i>to</i>) <i>miti</i>
ONE	<i>biðæ</i>	<i>ar</i>	<i>ena</i>
PATH	<i>yol</i>	<i>gza</i>	<i>orðomo</i>

	Tk. Qorveh	Laz	Rom.
PERSON	<i>kæs</i>	<i>k'otfi</i> (also ‘man’ and ‘husband’)	<i>nomat</i>
SEE	<i>görir</i>	<i>o-dzir-am-u</i>	<i>elepo</i>
SKIN	<i>daræ</i>	<i>tk'ebi</i>	
STAR	<i>oldüz</i>	<i>murutsxi</i>	<i>aspra</i> (PL)
STONE	<i>daf</i>	<i>kva</i>	<i>klersi</i>
SUN	<i>gün</i>	<i>mzora</i>	(to) <i>oulo</i>
TONGUE	<i>dil</i>	<i>nena</i> (organ)	(i) <i>ylossa</i>
TOOTH	<i>dif</i>	<i>k'ibiri</i>	(to) <i>ðond</i>
TREE	<i>Paqaſ</i>	<i>ndʒa</i>	<i>lisar, alat</i>
TWO	<i>ik'e</i>	<i>ʒur</i>	<i>ðio</i>
WATER	<i>su</i>	<i>tsk'ari</i>	<i>nero</i>
WATER-MELON	<i>qarpiz</i>	<i>k'arpuzi</i>	
WHY?	<i>niyæ</i>	<i>mot, mu-feni</i>	<i>oðen</i>
I	<i>mæn</i>	<i>ma</i>	<i>eyo</i>
WE	<i>biz</i>	<i>tʃku</i>	<i>emis(tin)</i>
YOU (SG)	<i>sæn</i>	<i>si</i>	<i>esi</i>
YOU (PL)	<i>siz</i>	<i>tkva</i>	<i>esis(tin)</i>

7.5. Notes on sources

Label	Variety, Sub-variety	Reference
Behd.	Northern Kurdish, Behdinī, spoken in Dohuk town-ship, Iraqi Kurdistan	Chap. 3.3, §4
	The Behdinī data were provided by Baydaa Mustafa, a female speaker from Dohuk. Due to difficulties in eliciting citation forms of verbs, all verb forms are given in the third person indicative present.	
Gor.I	Gorani, Šabakī, near Mosul, Iraq	Chap. 3.3, §5
	Gorani data were provided by Parvin Mahmoudveysi, who elicited them from a young male speaker from the village of Sofaya, between Erbil and Mosul, who claimed a Šabak identity. However, he was unable to provide all forms. Due to difficulties in eliciting citation forms of verbs, all verb forms are given in the third person masculine indicative present.	
Gor.II	Gorani, spoken in the town of Qorveh, Iran	
	Recorded by Masoud Mohammadirad in Qorveh, Iran, on 08. 03. 2018 and transcribed by Geoffrey Haig. The speaker is a retired male, 65 years old, from a Gorani-speaking family but married to a Central Kurdish speaking wife, and uses predominantly Central Kurdish in daily life. Verb forms are third person singular present. The transcription is broadly phonetic, using IPA symbols except for the following: the front rounded vowel [y] and the voiced palatal glide [j] are rendered with <ü, y> respectively, while all rhotics are rendered indiscriminately with <r>.	
Hawr.	Hawrāmī, spoken in the Hawrāmān region in the northern Zagros mountains in northwestern Iran and northeastern Iraq	Chap. 4.5
S. Kurd.	Southern Kurdish variety of Qorveh, Iran (see Fattah 2000 for Southern Kurdish)	
	The data were recorded in Qorveh, Iran, on 10. 03. 2018. The speaker is an educated male speaker (graduate student of linguistics, 29 years old), who has been living in France since September 2016. Verb forms are third person singular present. The transcription follows the one used for Gor.II (Gorani Qorveh).	
Bakht.I	Bakhtiari, spoken in Masjed Soleymān, Khuzestan Province, Iran (Source: Anonby and Asadi 2014)	Chap. 4.3
Bakht.II	Bakhtiari, spoken in Ardal, C&B Province, Iran	Chap. 4.3
Kumz.	Kumzari, spoken mainly in Kumzar and Khasab, Oman	Chap. 4.7

Label	Variety, Sub-variety	Reference
	<p>Although Kumzari is described as a mixed Semitic/Iranian language in van der Wal Anonby (chapter 4.7) we have included it provisionally under “Iranian” here, based on the fact that the majority of the lexical items of this list have Iranian etymologies (approx. 25); selected etymologies are the following: Semitic etymologies: SKIN: Arabic incl. Shihhi (Bernabela 2011:41); Mehri <i>ged</i> (Rubin 2010:17). TREE: Arabic; Mehri <i>śigarēt</i> (Watson 2012:25). HORN: Jibbali <i>kūhn</i> (Rubin 2014:37). KNEE: Arabic; Mehri <i>bark</i> (pl. <i>bīrōk</i>) (Rubin 2018:94). NOSE: Mehri <i>naxrīr</i> (Simeone-Senelle 1997:388); NENA I-II <i>naxira</i>. WATER-MELON: Kmz. <i>gahh</i>; Arabic <i>gīhh</i> (Holes 2001:83); Mehri <i>gawh</i> (Johnstone 1987:127). MOUNTAIN: Kmz. <i>aql</i>; Musandam Arabic <i>aql</i> (E.Anonby in prep.)</p> <p>Uncertain etymologies: SUN <i>intāf</i>: Mehri <i>ntaxf</i> ‘we come at sunset’. FROG <i>ṣufrāqō</i>: Mehri <i>ṣṣəfdēt</i> (Johnstone 1987:397). LOUSE <i>qar'a</i>: cf. NENA II <i>qalma</i>; Musandam Arabic <i>qamla</i> (E.Anonby in prep.). PATH <i>tēra</i>: Arabic <i>tariq</i>; Persian <i>rā</i>. BONE <i>xār</i>: cf. NENA <i>garma</i></p>	
Oss.	Ossetic, sub-dialects of Iron and Digor, spoken in North Ossetia-Alania, an autonomous republic within Russia, and South Ossetia, a breakaway part of Georgia	Chap. 6.3
	Iron and Digor forms are given simultaneously, they are shown in the following order and separated by a slash: <i>gədə/tikis</i> ‘cat’. Note on the dative of ‘why?’: this is etymologically the dative of <i>səči</i> ‘what’. Besides that, there exists a dedicated lexical item <i>səwənne/cəmənne</i> ‘why not’, etymologically <i>səwəl-ne</i> what.SUP-NEG and <i>cəmen-ne</i> what.DAT-NEG, respectively, see Erschler and Volk (2011) for arguments that these are single words. The personal pronouns are given for the nominative and oblique, other forms are as follows 1SG: Dat <i>menen</i> ; Abl <i>menej</i> ; All <i>menme</i> ; Sup <i>menəl/menbel</i> ; Equ <i>menaw</i> ; Com <i>menime</i> . 2SG: Dat <i>dəw-en</i> ; Abl <i>dəw-ej</i> ; All <i>dəw-me</i> ; Sup <i>dəw-əl/dəw-bel</i> ; Equ <i>dəw-aw</i> ; Com <i>dəw-imə</i> . 1PL: Dat <i>max-en</i> ; Abl <i>max-ej</i> ; All <i>max-me</i> ; Sup <i>max-əl/max-bel</i> ; Equ <i>max-aw</i> ; Com <i>max-imə</i> . 2PL: Dat <i>šəmax/sumax-en</i> ; Abl <i>šəmax/sumax-ej</i> ; All <i>šəmax/sumax-me</i> ; Sup <i>šəmax-əl/sumax-bel</i> ; Equ <i>šəmax/sumax-aw</i> ; Com <i>šəmax-imə</i> .	
Casp.I	Tatic (Lāhijāni), Eastern Gilaki	Chap. 5
Casp.II	Tatic (Sāravi), Mazanderani	Chap. 5
Tal.I	Tatic (Leriki), Northern Talyshi	Chap. 5
	Selected items (*) from Pirejko (1976) dictionary.	
Tal.II	Tatic (Māsulei)	Chap. 5
	Selected items (#) from Māsāli (Nawata 1982).	
Tat.I	Tatic (Koluri)	Chap. 5
Tat.II	Tatic (Vafsi)	Chap. 5

Label	Variety, Sub-variety	Reference
Arab.I	Arabic of Mardin (qəltu dialect)	Chap. 2.4
Arab.II	Arabic of Urfa/Harrān (Bedouin dialect)	Chap. 2.4
Arab.III	Mosul Arabic	Chap. 3.2
Arab.IV	Khuzestani Arabic, Iran The data were elicited from a native speaker consultant.	
CNA	Turoyo (Midin), Central Neo-Aramaic, spoken in the region of Tûr ‘Abdîn, southeastern Turkey (Source: Jastrow 1985 and native speaker consultants)	Chap. 2.5
NENA.I	Christian Urmî, North-Eastern Neo-Aramaic, spoken on the plain of Urmî and within the town of Urmî, Iran (Source: Khan 2016)	Chap. 2.5
NENA.II	Christian Barwar, North-Eastern Neo-Aramaic, spoken in the Barwar region of northern Iraq close to the Turkish border, in a number of villages along the Be-Xelape River. (Source: Khan 2008)	Chap. 3.4
NENA.III	Jewish Sanandaj, North-Eastern Neo-Aramaic, spoken in western Iran (Source: Khan 2009)	Chap. 4.4
NM	Khorramshahr, Neo-Mandaic, spoken until recently in the cities of Ahwaz and Khorramshahr, Iran (Source: Häberl 2009; Mutzafi 2014)	Chap. 4.4
Tk. Qorveh	Iran Turkic, dialect of the town of Qorveh, Iran Recorded in Qorveh, Iran, on 08. 03. 2018, by Geoffrey Haig. The speaker is an educated male (high school teacher, approx. 30 years old) raised in a primarily Turkish speaking family in Qorveh. He is trilingual, regularly speaking the local varieties of Turkish and a mixture of Central and Southern Kurdish, and Persian, plus a little Gorani. His mother (approx. 60 years old, primarily Turkic speaking) was also present at the recording. Verb forms are third person singular present. A phonemic analysis was not possible based on this minimal data set, but see Bulut (this volume, chapter 4.2) for a description of the relatively close dialect of Sonqor. The transcription is broadly phonetic, using IPA symbols except for the following: the two front rounded vowels [ø, y] and the voiced palatal glide [j] are rendered with <ö, ü, y> respectively, while all rhotics are rendered indiscriminately with <r>.	
Laz	Laz (Kartvelian), spoken in northeastern Turkey near the Black Sea, and in Georgia, mainly in the village of Sarp. The lexical list is based on the variety of Arhavi in Turkey.	Chap. 6.2

Label	Variety, Sub-variety	Reference
Rom.	Romeyka, spoken in north-eastern Turkey in Trabzon Province, data based on the variety of Çaykara (ROf)	Chap. 6.4
For nouns, the article is provided where available. Due to the lack of an independent infinitive form, verbs are provided in the 1st person singular which is consensus in Greek philology.		

References

- Anonby, Erik. In preparation. *Dialectology of Musandam Arabic*.
- Anonby, Erik & Ashraf Asadi. 2014. *Bakhtiari studies: phonology, text, lexicon*. (Studia Iranica Upsaliensia 24). Uppsala: Acta Universitatis Upsaliensis.
- Erschler, David & Vitaly Volk. 2011. On negation, negative concord, and negative imperatives in Digor Ossetic. In Geoffrey Haig, Simin Karimi, Agnes Korn & Pollet Samvelian (eds.), *Topics in Iranian Linguistics* (Beiträge zur Iranistik Bd. 34), 135–152. Wiesbaden: Reichert.
- Fattah, Ismail Kamandâr. 2000. *Les dialectes kurdes méridionaux: étude linguistique et dialectologique*. (Acta Iranica 37). Leuven: Peeters.
- Holes, Clive. 2001. *Dialect, culture, and society in Eastern Arabia: Glossary*. Leiden & Boston: Brill.
- Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List & Dmitry Egorov. 2011. Automated dating of the world's language families based on lexical similarity. *Current Anthropology* 52(6). 841–862.
- Häberl, Charles. 2009. *The Neo-Mandaic dialect of Khorramshahr*. (Semitica Viva 45). Wiesbaden: Harrassowitz.
- Jastrow, Otto. 1985. *Laut- und Formenlehre des neuaramäischen Dialekts von Mīdin im Tūr ‘Abdīn*. Wiesbaden: Harrassowitz.
- Johnstone, Thomas M. 1987. *Mehri lexicon and English-Mehri word-list*. London & New York: Routledge.
- Khan, Geoffrey. 2008. *The Neo-Aramaic dialect of Barwar*. 3 vols. Leiden & Boston: Brill.
- Khan, Geoffrey. 2009. *The Jewish Neo-Aramaic dialect of Sanandaj*. Piscataway, NJ: Gorgias Press.
- Khan, Geoffrey. 2016. *The Neo-Aramaic dialect of the Assyrian Christians of Urmi*. 4 vols. Leiden & Boston: Brill.
- Mutzafi, Hezy. 2014. *Comparative lexical studies in Neo-Mandaic*. (Studies in Semitic Languages and Linguistics 73). Leiden & Boston: Brill.
- Nawata, Tatsuo. 1982. The Masal dialect of Talishi. *Acta Iranica* 22. 93–117.
- Pirejko, Lija A. 1976. *Talyšsko-russkij i russko-talyšskij slovar* [Talyshi-Russian and Russian-Talyshi dictionary]. Moskva: Izdatel'stvo Russkij jazyk.
- Rubin, Aaron D. 2010. *The Mehri language of Oman*. Leiden & Boston: Brill.
- Rubin, Aaron D. 2018. *Omani Mehri: A new grammar with texts*. Leiden & Boston: Brill.
- Simeone-Senelle, Marie-Claude. 1997. The modern South Arabian languages. In Robert Hetzron (ed.), *The Semitic languages*, 378–423. London & New York: Routledge.
- Watson, Janet C.E. 2012. *The structure of Mehri*. Wiesbaden: Harrassowitz.
- Wichmann, Søren, Eric W. Holman & Cecil H. Brown (eds.). 2016. *The ASJP Database* (version 17).

Index of language names

The index includes names of languages, language groups, varieties, and dialects.

A

- Abivardi [Abiverdi] (variety of Qashqâ'i) 417, 434
Adiyaman-Urfâ, Kurmanjî of 126
Afshâr [Afshar] dialects → Qashqâ'i (Central Oghuz)
Akkadian 237, 305, 309
Akre, Kurmanjî of 288–291, 294
Alvîri 672, 678, 689, 694, 700, 711, 712, 717, 725–727, 752, 754, 767, 770, 772, 820, 823
Amêdî, [Amêdi, Amedia] Kurmanjî of 16, 36, 37, 137, 306, 312–315, 317–319, 322, 486
Âmoli 719, 824
Anbarâni [Anbarani] 660, 712, 824
Ankawa, Christian NENA 11, 239, 306, 313, 323, 324, 337, 350
Anzali 659
Arabian, South 7, 627–629, 638, 639, 642, 658, 951
Arabic 3, 7, 8, 10, 11, 13, 16, 19, 20, 22, 23, 28, 31, 32, 35–44, 77, 108, 109, 150, 159–177, 179–184, 187–189, 196, 198, 226–228, 237, 239, 240, 243–250, 252, 254, 259, 260, 263–266, 270–272, 288, 290, 296, 306, 311, 316, 325, 341, 344, 345, 351, 353, 357, 358, 360–363, 369, 372, 377, 386, 387, 406, 408, 409, 411, 413, 414, 417, 427, 440–442, 450, 484, 485, 490, 492, 495, 524, 526, 541, 570–574, 579, 581, 583, 585–588, 607, 624, 627–629, 636, 638, 649, 656, 657, 677, 686, 692, 827, 895, 948, 949, 951
Arabic, Anatolian 13, 16, 20, 35, 159, 163, 166, 187, 188, 189
Arabic, bedouin-type dialects 31, 42, 159, 164, 181, 184, 188, 196, 243, 244, 252, 259, 629, 949
Arabic, dialects of 10, 11, 13, 16, 23, 28, 31, 32, 35–37, 39, 41, 42, 159, 161, 179, 181, 183, 187, 188, 228, 239, 240, 243, 245, 259, 263–266, 316, 345, 492, 629, 649, 656, 657, 827
Arabic, Harran-Urfâ 162–171, 173–175, 177, 179–184
Arabic, Jewish Arabic dialects 259, 265
Arabic, Mardin 164–165, 170–171, 177, 949
Arabic, Mesopotamian 45, 187, 243, 264, 265
Arabic, Northern Iraqi (vernacular) 160, 246, 249, 260, 827
Arabic, *qəltu* [qeltu] dialects 31, 32, 35, 159, 160, 162, 164–168, 170–173, 175, 176, 179–183, 187, 244, 248, 249, 263–265, 949
Arabic, Anatolian *qəltu* dialects 160, 165–167, 172, 175, 176, 179, 180, 188, 265
Arabic, Šâwî [Šâwi] 159, 188
Arabic, sedentary-type dialects 31, 159, 244, 252, 259, 260
Arabic, Syrian 163, 170
Aradhin, Christian NENA 306, 324
Aramaic 9, 14, 35–37, 160, 162, 167, 173, 183, 197, 201, 213, 214, 219, 220, 225–228, 234–236, 238, 243, 254, 260, 266, 305, 309, 311, 312, 316, 326, 328–330, 341, 342, 344, 350, 352, 353, 386, 394, 485, 488, 489, 501, 503, 505, 513, 522–524, 526, 532, 625, 828, 843
Aramaic, Jewish Babylonian 305
Aramaic, Middle 350, 513
Aramaic, Central Neo-Aramaic → Central Neo-Aramaic (CNA)
Aramaic, North-Eastern Neo-Aramaic → North-Eastern Neo-Aramaic (NENA)
Arbel, Jewish NENA 240, 306, 312, 314, 322–327, 335, 337, 339, 341, 343, 345, 351
Armenian 1, 5, 8–10, 14, 16, 18, 19, 21, 24, 25, 27, 31–34, 36, 37, 39–43, 45–50, 65, 66, 68, 70, 72–76, 79, 81, 83, 85, 87, 89, 90, 92–104, 108, 109, 113, 121, 150, 151, 196, 201, 218, 238, 336, 352, 357, 406, 748, 786, 790, 840, 895
Asâlemi 671, 672, 677, 685, 686, 691, 694, 698, 702, 712, 713, 715, 716, 718, 723, 724, 731, 736, 737, 739, 743–745, 748, 756, 765–767, 780, 782, 785, 787, 791, 793, 824
Ashitha, Christian NENA 33, 35, 39, 43, 44, 197, 198, 201–206, 210, 213, 215, 217, 218, 223, 224, 226, 227, 233
Astara 660, 664
Avestan 46, 571
Azd 625, 627
Azerbaijani (Azerbaijan Turkic, Central Oghuz) 376, 398, 399, 407, 416, 423–425, 444, 621, 663, 763, 822, see also Azeri Turkic

- Azeri Turkic 12, 33, 196, 200, 228, 229, 259, 355, 364, 365, 367, 376, 377, 398–400, 406, 408, 412, 421, 439, 444, 570, 577, 578, 673, 676, 679, 763, 806, see also Azerbaijani
- B**
- Bābolsari 678, 686, 687, 719, 757, 824
- Bājałānī [Bājālānī, Bājalānī, Bājalānī] 297, 303
- Bakhtiari [Bakhtiāri] 12, 25, 385–389, 391–393, 395, 419, 445–480, 629, 947, 951
- Bākolumi 664
- Balochi [Baluchi] 6, 25, 627
- Barwar, Christian NENA 16, 20, 45, 157, 238–240, 306, 312–315, 317–320, 322, 324–327, 330, 332–334, 337–340, 342–344, 346, 351, 500, 501, 949, 951
- Barətlı, Christian NENA 306, 317
- Bañnaya, Christian NENA 323, 324
- Bayat (variety of Qashqā’i) 416, 422–425, 430
- Baz, Christian NENA 199, 204, 205, 209, 210, 212, 638
- Baz, Aruntus, Christian NENA 205, 210
- Baz, Maha Xtaya, Christian NENA 204, 209, 210, 212
- Begdilli [Begdili] (variety of Qashqā’i) 423–425
- Behdinī [Bahdini, Badini, Badinī, Bahdinī, Behdini], Northern Kurdish 13, 24, 34, 40, 41, 43, 44, 106, 108, 114, 119, 121, 125–127, 129, 132, 134, 137, 138, 145, 267, 269, 272, 280, 287, 292, 294, 304, 377, 388, 947
- Bēşpən, Christian NENA 198
- Betanure, Jewish NENA 306, 310, 312, 314, 323, 324, 336, 352
- Bēwyānī (variety of Gorani) 297
- Bidaro, Christian NENA 313, 318
- Bingird, Sorani of 271, 285
- Bohtan, cluster of Christian NENA 14, 16, 32, 34, 35, 40–42, 191, 194, 198–205, 207, 210, 211, 216–218, 220, 221, 224, 225, 228, 233, 234
- Bokan, Jewish NENA 481, 496, 508
- Bədylə, Christian NENA 310, 313, 314, 318, 319, 324, 335, 337, 339
- C**
- Cappadocian 893, 901, 903, 932
- Caucasian, South 9, 830, 843, 846, 854, 855, 860, 885, 889
- Caucasian Tat 659, 664, 691, 819, 820, 823, 578, 938–940, 948
- Central Neo-Aramaic (CNA) 6, 10, 14, 16, 17, 20, 31, 32, 34–37, 41, 42, 190, 194, 197, 198, 200–209, 211–213, 215, 219, 227, 228, 234, 941–944, 949
- Chālı 664, 672–675, 678, 685, 687, 694, 696, 711, 715, 721, 733, 739, 767, 775
- Charazei 664
- Charozh 664, 716, 737, 824
- Chenār Rudxāni 664
- Chikolen-Lesken [Chikola-Lesken] dialect of Digor Ossetic 862
- Chinese 574, 601
- Cudi cluster of Christian NENA 191, 198, 204, 205, 210, 224
- D**
- Darī [Dari] 569, 570, 572, 579, 603–605, 607, 609, 610, 612, 623
- Derabun, Christian NENA 315
- Dere, Christian NENA 324
- Digor Ossetic 861–863, 865–874, 876, 877, 886, 887, 889, 890, 948, 951
- Dikin-Marāqeи 665, 678, 717, 727, 735, 780, 782, 824
- Dikin-Pashei [Dikin-Pašeи] 665, 674, 675, 824
- Diyana-Zariwaw [Diyana-Zariway] 312, 313–318, 322, 330, 352
- Dīz, Christian NENA 203
- Dizi 664, 730, 731, 824
- Dobe, Jewish NENA 323, 337
- Dohuk, Jewish NENA of 306, 314, 323
- Dohuk, Kurmanjî of 138, 290, 291, 303
- Domari 3, 28, 29
- E**
- Elästi [Elašti] 666, 668–670
- Elazığ [Elaziğ], Kurmanjî of 127
- Elbistan, Kurmanjî of 121, 127, 136, 139, 141, 149
- Erzurum, Kurmanjî of 127
- Eştehārdi [Eştehārdi] 724, 738, 780
- F**
- Fârsî [Farsi, Fârsî] 405, 406, 434, 485, 569, 570, 573, 579, 580, 582, 589, 600–602, 605, 607, 609, 610, 612, 616, 620–624, 821
- Fumanı 659
- G**
- Galeshi [Galeši, Gâleši] 663, 666, 668–670, 675
- Gandomäbi 664, 678, 723, 724, 730, 731, 736, 738, 759, 772, 778, 824
- Gawrajüyî [Gawfajüyî] (variety of Gorani) 296, 533, 541

Gāzарxāni (Tatoid) 674, 675
 Georgian 9, 10, 196, 748, 763, 825, 826, 830, 833, 840, 855, 862, 878, 885, 891, 895
 Geylavāni 664
 Gīlakī [Gilaki] 18, 19, 577, 578, 658–660, 665, 673–677, 679, 688, 691, 692, 697, 709, 716–718, 729, 769, 820, 821, 948
 Gorani [Gürānī, Gūrānī, Gurānī, Gurāni] 7, 13, 15, 237, 267, 269, 270, 274, 278, 281, 295–302–304, 357, 368, 385, 410, 413, 426, 533, 534, 536, 537, 540, 566, 567, 947, 949
 Gulf Arabic 629, 657
 Gullī, Kurmanjī of 291
 Gulpašan, Christian NENA 200
 Gurāni (variety of Tatoid) 674, 675, 700, 703, 714, 728, 735, 771, 824
 Gurchāni (Vafsi) 665, 676

H

Hakkari cluster of Christian NENA 191, 199, 201, 203–206, 209, 210, 212
 Halmun, Christian NENA 204, 224
 Hamziye, Christian NENA 324
 Harbole, Christian NENA 205
 Harzani [Harzandi] 40, 664, 671, 672, 676, 689, 698, 710–712, 718, 720, 728, 731, 733, 739, 744, 749–751, 764, 770, 771, 791–793, 820, 821, 823
 Hashtpari 660, 820
 Hassana, Christian NENA 198, 210
 Hawrami [Hawrāmī, Hawrāmī] 10, 15, 18, 25, 267, 295, 296, 298, 299, 301, 385–394, 533–539, 541–543, 546, 553, 554, 566, 567, 763, 947
 Hertevin, Christian NENA 14, 16, 34, 35, 198, 200, 201, 203, 205, 210, 211, 216–218, 221, 234, 526
 Hezārrudi 664, 678, 695, 717, 730, 731, 736, 767–769, 778, 779, 785, 800, 801
 Hoboyot 629

I

Indo-Aryan 639
 Indo-European 53–60, 104
 Indo-Iranian 625, 627
 Iranian languages 3, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 18, 20, 21, 22, 24, 25, 26, 28, 29, 40, 45, 94, 106, 108, 118, 124, 130, 151, 156, 157, 158, 189, 195, 214, 227, 229, 237, 238, 239, 240, 241, 260, 267, 269, 272, 273, 274, 276, 278, 280, 285, 287, 295, 296, 298, 299, 303, 304, 306, 327, 354, 362, 363, 364, 366, 368,

369, 370, 371, 374, 376, 383, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 398, 400, 401, 411, 413, 414, 415, 416, 419, 420, 423, 426, 427, 428, 432, 433, 434, 435, 436, 437, 438, 439, 444, 445, 447, 458, 460, 472, 473, 478, 479, 480, 484, 485, 494, 503, 522, 523, 533, 538, 566, 567, 568, 569, 570, 571, 572, 573, 575, 576, 577, 578, 579, 601, 621, 622, 623, 624, 625, 627, 629, 638, 639, 657, 658, 659, 660, 664, 667, 708, 711, 716, 720, 730, 731, 763, 778, 790, 820, 821, 822, 823, 825, 826, 827, 828, 829, 861, 884, 885, 889, 890, 936, 948, 951
 Iranian, Western 124, 385, 494, 623–625, 629, 639, 658, 821
 Iranic → Iranian
 Iraq Turkic → Turkic, Iraq
 Iron Ossetic 861, 862, 865, 889
 Iron Ossetic, Kudar dialect 863, 878

J

Jāfi 534, 537, 538
 Jamshidābādi 673
 Jilu, Christian NENA 11, 199, 203, 206, 233, 234
 Jowkandāni 712, 727, 728, 824
 Jubani 673–675, 700, 739, 824

K

Kabatei [Kafteji] 664, 665, 671–675, 678, 685, 694, 696, 705, 711, 712, 717, 727, 728, 737, 755–757, 778, 779, 822, 824
 Kāboli [Kaboli] 610
 Kajali 664, 671, 672, 678, 717, 723, 737, 738, 764, 767–769, 773, 774, 797, 824
 Kaka'yī [Kākayī] (variety of Gorani) 297, 533
 Kalāsari [Kalāsuri] 40, 664, 671, 685, 694, 703, 712, 718, 721, 725–728, 731, 733, 764, 768, 769, 823, 824
 Kalhor of Bayâdestân 424, 435, 436
 Kalhorī 577
 Kandelusi 666
 Kandūlayī [Kāndūlāī] (variety of Gorani) 296, 444, 533, 5, 567
 Karamlesh, Christian NENA 323, 324
 Karani 664, 671, 672, 707, 723–725, 736, 738, 823, 824
 Kārgānrudi [Kārgānrudi] 764, 766
 Karnaqī 664, 730, 731, 738, 824
 Kartvelian 5, 8, 10, 21, 29, 241, 401, 825–827, 830, 860, 949
 Kelārdasti [Kelārdashti] 660, 666, 673, 674, 675, 692, 693, 718, 720, 721, 729–731, 735

- Keläsi [Kelasij] 664, 665, 673–675, 678, 685, 687, 694, 703, 715, 727, 736, 738, 752, 754, 757, 822, 824
- Kerend, Jewish NENA 481, 496, 497, 503, 508, 512, 532
- Keringāni [Keringani] 664, 671, 710–713, 718, 720, 721, 725–727, 730, 731, 744, 751, 758, 764, 791
- Khalaj 19, 398, 399, 405, 407, 411, 415, 416, 419, 422, 424, 430, 431, 433, 434, 438, 439
- Khalaj of Bayâdestân 19, 407, 415, 422, 424, 430, 431, 433, 434, 438, 439
- Khatirâbâdi 824
- Khoini [Xo’ini] 694, 698, 717, 723, 727, 728, 730, 731, 733, 739, 753, 823, 824
- Khorâsân-Turkic (East Oghuz) 398
- Khoynarudi 664, 711, 712, 733, 774, 784, 787, 788, 800, 801, 803, 824
- Kilit 659, 664
- Koluri 664, 671, 672, 677, 678, 685, 687, 691, 698, 711, 715–717, 720, 723, 724, 727, 728, 730, 731, 736, 748, 819, 824, 948
- Kordkheyli 678, 686, 697, 747, 808, 809
- Koy Sanjak [Koy Sanjaq, Koy Sanjaq], Jewish NENA 296–298, 305, 306, 312, 313, 315, 316, 323, 352
- Kumzari [Kumzârî, Komzari, Kamazara] 1, 7, 15, 22, 24, 25, 385, 387, 389, 390, 395, 458, 480, 571, 625, 627–636, 638–651, 653, 656, 658, 947, 948
- Kurdish 3, 6–18, 20, 24, 25, 28, 31–35, 37–39, 41, 44, 45, 106–110, 113, 115–119, 121, 122, 126, 129–131, 134, 135, 137, 139, 143, 148, 150–152, 156–158, 160, 162, 163, 177, 181–184, 188, 189, 194, 196, 200, 204, 213, 226–228, 237–240, 243, 244, 246, 251, 259, 260, 267, 269–274, 276, 278, 280, 281, 284, 287–291, 293, 294, 296, 298–304, 310, 311, 314, 316, 321–323, 325, 332, 337, 341, 344, 345, 351, 352, 354, 355, 357–361, 368–370, 374, 375, 377, 383, 388, 389, 392, 400, 406, 407, 410, 412, 413, 426–428, 434, 437, 441, 445, 460, 468, 482, 485, 490, 501, 509, 526, 532–534, 536–538, 553, 561, 566, 571, 577, 601, 665, 699, 705, 826, 827, 837, 840, 843, 895, 918, 947, 949
- Kurdish, Northern (Kurmanjî) 8, 10, 13, 16, 17, 24, 25, 106, 131, 150, 156, 238, 267, 270, 271, 273, 278, 280, 284, 287–289, 291, 298, 370, 388, 468, 705, 947, see also Kurmanjî
- Kurdish, Central (Sorani) 10, 11, 18, 25, 41, 137, 139, 143, 213, 237–241, 267, 269–274, 276–278, 280–285, 287, 288, 291–293, 298, 299, 302, 304, 316, 322, 325, 358, 369, 370, 485, 501, 534, 536, 537, 566, 826, 827, 947
- Kurdish, Southern 3, 15, 106, 274, 278, 281, 296, 298–300, 368, 460, 947, 949
- Kurmanjî [Kurmanji, Kurmandji, Kurmanci, Kurmandjî Kurmângî] 10, 29, 31, 34–37, 39–45, 106–111, 113, 115, 117, 118, 120–122, 124, 126, 127, 130, 131, 133–139, 141, 144, 145, 147–152, 156–158, 184, 189, 237–240, 267, 270, 281, 287, 289, 291, 303, 314, 316, 369, 427, 673, 699, 837, 840, 918, see also Kurdish, Northern
- Kurmanjî, Southeastern (SEK) 40, 42–44, 133–137, 139, 140, 142–146, 148, 150, 151
- Kurmanjî, Southern (SK) 133–136, 139, 144, 148, 149
- Kurmanjî, Western (WK) 108, 126, 127, 131, 133–136, 138, 139, 141, 145–151
- L**
- Lâhijâni [Lahijâni, Lâhijani] 19, 659, 660, 666, 667, 669, 670, 677–679, 682, 685, 686, 688, 691–693, 697, 701, 703–705, 708, 709, 713, 715, 718, 719, 720, 721, 723, 725–727, 729–732, 734, 735, 737, 738, 741–745, 747, 749–752, 754, 757–763, 765, 766, 768, 769, 771, 792, 793, 795–803, 805, 808, 809, 824, 948
- Lâkei [Lakei] 673
- Laki [Lakî] 300, 303, 577
- Langerudi 659, 666, 734, 735, 743, 745, 749, 824
- Lârestânî 658
- Latin 46, 163, 357, 574, 676
- Latin, Low 574
- Laz 8, 10, 13, 16, 19, 21, 24–26, 241, 825–828, 830, 832, 833, 835, 837–841, 843, 844, 846, 847, 850, 852–854, 856, 860, 895, 923, 945, 946, 949
- Laz, dialect of Arhavi-Findikli 832, 833, 835, 844, 852, 854, 860, 949
- Laz, dialect of Hopa 72
- Lerdi 664, 733, 824
- Leriki 660, 663, 677–679, 683–687, 690, 694, 695, 698, 702–704, 706, 708–714, 716, 719, 722–728, 731, 732, 734, 739–741, 743–746, 748, 751–760, 763–769, 772, 774, 780, 781, 785, 786, 790, 791, 793, 797–799, 801–809, 814, 824, 948
- Levandavili 663
- lišana deni Jewish NENA 191, 306

- Lorī [Luri, Lori] 407, 445, 446, 452, 478, 479, 571, 629
 Lori, Northern 445
 Lori [Luri], Southern 445, 452, 629
 Lənkəran [Lenkoran] dialect 660, 663, 824
- M**
- Mäčiāni 670
 Mačo [Mäčo] (variety of Gorani) 297–299, 301, 534
 Mandaic 305, 393, 397, 485, 488, 502, 526, 532
 → Neo-Mandaic
 Mandaic, Classical 393, 485
 Mangesh, Christian NENA 324
 Maras̄ [Maras̄], Kurmanjî of 52, 156
 Mardin, Kurmanjî of 34, 116, 139, 141, 145, 152, 159, 160
 Mar-Yaqo, Christian NENA 324
 Mäsāli 671, 672, 675, 685, 824, 948
 Masally 659, 660, 663
 Maşlāwi Arabic → Mosul, Arabic of
 Mäsulei 660, 671, 672, 677, 678, 681, 685, 694, 703, 710, 715–717, 720, 721, 723, 725, 731, 736, 737, 739, 741, 744–748, 756, 758–763, 765–767, 769, 770, 772, 773, 775, 780, 793, 795, 799, 807, 809, 824, 948
 *Mawana, Christian NENA 13, 34, 202, 204, 213
 Mäzandarānī [Mazanderani, Mäzanderänī, Mazandarani] 577, 624, 659, 660, 677, 678, 691–693, 697, 709, 714, 715, 719, 735, 741, 747, 762, 821, 822, 948
 Mazra, Christian NENA 218
 Mehri 627, 629, 638, 658, 948, 951
 Middle Persian → Persian, Middle
 Midyat, Kurmanjî of 107, 133, 144, 148, 152
 Mingrelian 830, 885, 890
 Mlahsō [Mlahso] 32, 41, 42, 45, 190, 195, 197, 198, 204–207, 209, 213, 220, 221, 234, 526
 Mongolian 4, 574
 Moşul, Arabic of 23, 240, 246, 247, 248, 250, 252, 260–265, 949
 Mukri, Sorani of 12, 34, 270–272, 282, 304
 Muş [Muş], Kurmanjî of 8, 132, 159
- N**
- Nargezine-Xarjawa, Christian NENA 323
 Nāvrudi 663
 NENA → North-Eastern Neo-Aramaic
 Neo-Mandaic 19, 21, 22, 385–388, 390, 393–395, 481–485, 487–493, 495, 497, 498, 500–502, 504, 513–515, 517, 519, 521, 522, 524, 526, 529, 531, 532, 949, 951

- Nerwa, Christian NENA 11, 306, 312, 313, 353
 Nerwar, Jewish NENA 310, 314, 323, 324
 North-Eastern Neo-Aramaic (NENA) 6, 8–17, 19–23, 26–28, 31–44, 157, 190, 191, 194–207, 209–228, 230, 235, 238, 241, 269, 292, 305–307, 309–317, 319–322, 324, 328–334, 336, 345, 350, 351, 385–397, 460, 481–486, 488–494, 496–505, 507–513, 515, 516, 518, 521, 522, 525–527, 531, 532, 826, 941–944, 946, 948, 949
 NENA, Christian (dialects) 9, 13, 191, 194, 196, 198, 203, 206, 215, 222, 223, 306, 308, 311, 388, 390–392, 482, 484, 486, 488, 489, 491
 NENA, Jewish (dialects) 194, 199, 241, 269, 306, 307, 311, 316, 387, 388, 390–394, 481, 488, 491, 507, 508, 511, 512, see also trans-Zab Jewish NENA
 Nowkiāni [Nowkiani] 664, 671, 678, 689, 720, 721, 725, 727, 728, 778, 784, 788, 824
- O**
- Oghuz, Central (Turkic) 358, 399, 439
 Oghuz, South (Turkic) 399, 405
 Ophitic 894, 932
 Oskulaki 665, 673
 Ossetic 8, 12, 16, 24–26, 825–829, 861–863, 865–868, 871, 873, 876–878, 883–891, 948, 951
- P**
- Pahlavī [Pahlavi] 572, 623
 Pārsī [Parsi, Pārsig] 572, 622, 623
 Parthian 274, 571, 572
 Pashto 6, 570, 577, 579
 Pāwayī 536
 Persian 3, 18, 22, 24, 25, 27, 65, 77, 85, 87, 118, 176, 196, 228, 229, 234, 243, 274, 300, 303, 362, 366, 368–370, 375, 377, 385–387, 389–393, 395, 396, 401, 405–409, 411–415, 418, 424–426, 428, 429, 433, 434, 436–442, 445, 447, 449, 450, 458, 460, 464, 472, 473, 477, 479, 480, 484, 485, 490, 501, 502, 504, 509, 514, 520, 524–526, 533, 534, 537, 538, 569–593, 595–597, 600–604, 607, 608, 614, 617, 621–625, 627–629, 637, 657, 658, 664, 665, 667, 668, 673, 677, 679, 687–689, 691, 692, 697, 702, 704, 717, 735, 738, 756, 757, 763, 786, 796, 798, 799, 806, 820, 821, 840, 855, 884, 885, 915, 948, 949, see also Fārsī
 Persian, Classical 576, 578, 582, 592, 623

- Persian, Colloquial 24, 27, 370, 389, 390, 437, 538, 576, 578, 583, 586–588, 590, 592, 596, 614, 616, 620, 621, 624
- Persian, Early Judaeo- 392, 580, 624
- Persian, Early New 576, 580, 582, 584, 624
- Persian, Judaeo- 392, 572, 580, 624
- Persian, Middle (MP) 274, 571–573, 584, 624, 627, 637
- Persian, Middle, Manichaean 584
- Persian, Middle, Zoroastrian 572, 584
- Persian, Modern 368–370, 401, 405, 407, 408, 409, 411, 413, 436, 479, 588, 593, 604
- Persian, New (NP) 571–573, 576, 580, 582–585, 590, 596, 602, 624
- Persian, Old 571
- Persian, Standard 385, 390, 538, 576, 580, 581, 583, 586, 588, 589, 608
- Peshabur, Christian NENA 313, 322, 324
- Piždar, Sorani of 271, 285
- Pontic Greek 893–895, 897, 903, 907, 932, 933
- Pontic Greek, Muslim → Romeyka
- Q**
- Qal’ei 824
- Qar Hasan, Jewish NENA 508
- Qaraqosh, Christian NENA 10, 11, 240, 306, 309, 311–313, 315–324, 326, 328, 329, 333, 334, 337, 340–345, 351
- Qashqā’ī (South Oghuz) 368, 383, 399, 405, 407, 409–414, 416–418, 421, 423–426, 430, 444
- Qočanəş, Christian NENA 201
- R**
- Rāmsari [Ramsari] 659, 666, 669, 670, 688, 692, 693, 709, 721, 730–732, 734, 735, 743, 745, 747–749, 757, 821, 824
- Rashti [Rašti] 666, 668–670, 673–675, 677–679, 685, 688, 697, 711, 717, 720, 725–727, 730, 731, 734, 736, 738, 741, 751, 762, 765, 824
- Razajerdi 664, 678, 689, 694, 700, 703, 720, 727, 729, 739, 747, 770, 771, 786, 791, 824
- Rewandiz, Sorani of 292
- Romeyka 12, 13, 16, 20, 21, 825–828, 892–895, 897–907, 909–912, 914–927, 929–931, 933, 950
- Rostamābādi 673, 824
- Rōžbayānī [Rožbayānī] (variety of Gorani) 533
- Rudbāri 659, 665, 667, 673–675, 700, 716, 717, 720, 727, 728, 730, 731, 734, 739, 762, 771, 824
- Ruma, Christian NENA 14, 34, 42, 191, 203, 207, 216–218, 228
- Rumca → Romeyka
- Rumdža → Romeyka
- Rumta, Christian NENA 201
- Russian 45, 100, 194, 196, 570, 575, 861, 863, 886, 887, 891
- Rustaqa, Jewish NENA 306, 323, 335–337, 351
- S**
- Šabakī (variety of Gorani) 296, 298–300, 533, 947
- Saff Khānī (variety of Qashqā’ī) 416, 417, 423–425, 430
- Sagzābādi 664, 676, 740, 785, 788, 801, 804
- Saljuqian Turkic (historical form) 398–400
- Šammar dialects 243, 244, 246
- Šammāsdin-Gawar cluster of Christian NENA 191, 202–204, 213
- Sanandaj, Christian NENA 389–392, 482, 484, 485, 487–494, 497, 497, 499, 501, 503–505, 509–512, 516–519, 521–526
- Sanandaj, Jewish NENA 10, 15, 19–21, 23, 351, 386, 387, 389, 393, 394, 397, 481, 482, 485–492, 494–497, 499–501, 503, 506–512, 517–519, 521–524, 527, 532, 949, 951
- Sangesari 280
- Saqəz, Jewish NENA 481, 494, 497
- Sāravī 660, 666, 668–670, 677, 678, 683, 685–687, 692, 693, 697, 698, 701, 703–705, 707–709, 711, 715, 719, 721, 727, 730–732, 735, 736, 741–743, 745, 747, 749–751, 753, 757–759, 762, 763, 765, 766, 768, 769, 771, 792–795, 797–799, 801, 804, 808, 809, 812, 948
- Sarlı 267
- Sayyādları 824
- Semitic 3–7, 10, 14, 20, 21, 25, 27, 29, 31, 32, 42, 113, 150, 159, 189, 234–241, 243, 254, 264–266, 272, 292, 305, 328, 350, 351, 353, 361–363, 366, 371, 385, 387, 388, 397, 401, 410, 411, 531, 532, 625, 627, 630, 632, 635, 643, 658, 763, 821, 822, 825–827, 829, 941, 948, 951
- Semitic, South 625
- Semnānī [Semnani] 571, 577
- Šēxānī [Şēxānī] (variety of Gorani) 298, 533
- Šēx Bizimi 106
- Shahrāni 673
- Šāli 664
- Šāndermeni 664
- Shihhi [Shihhuh, Shihuh] 627–629, 656, 658, 948
- Šōsh-u-Sharmə, Christian NENA 323
- Siāvarudi 664

- Sinjar, Kurmanjî of [Sinjarî] 288, 290
 Sorani → Kurdish, Central
 Sulemaniyya, Christian NENA 305, 313, 315, 323, 324, 333, 487
 Sulemaniyya, Jewish NENA 15, 241, 312, 319, 321, 323, 324, 327, 329–332, 334, 335, 339–342, 344, 345, 482, 505, 510
 Sulemaniyya, Sorani of 270–273, 277–279
 Sumerian 237
 Surči [Surči], Kurmanjî of 281, 291
 Svan 830, 885, 891
 Syriac 196, 208, 219, 305, 310, 311, 351, 352, 489
 Šəno, Jewish NENA 191, 213
- T**
- Tabarî (New Tabarî) 572, 621
 Tahāromi 664
 Tajik [Tojiki Tâjikî] 458, 569, 570, 579, 580, 582, 593, 596, 597, 602–607, 609–612, 614, 616, 617, 624
 Täkestäni 664, 675
 Täleš-Dolabi 664
 Täleši [Täleshî, Taleshi] 571, 578, 659, 819–821, 821, 823, 951
 Telkepe, Christian NENA 318, 323, 324, 334, 350, 352
 Tikab, Jewish NENA 508
 Tisqopa, Christian NENA 323
 Tiyare cluster of Christian NENA 191, 197–199, 201–206, 210, 213, 215, 224, 226–228, 306
 Tiyare, Lower, Christian NENA 191, 198
 Tiyare, Upper, Christian NENA 191, 199, 218
 Tonekâbonî [Tonekâbonî] 677, 688, 692, 693, 717–719, 721, 727, 730–732, 734, 743, 745, 749, 757, 762, 807, 819, 824
 trans-Zab Jewish NENA 21, 26, 191, 199, 213, 222, 223, 235, 238, 239, 241, 306, 314, 319, 321, 322, 326, 327, 339, 345, 352, 481, 486, 492, 516, 518, see also NENA, Jewish
 Tur Abdin [Tûr ‘Abdîn, Tur ‘Abdîn], Kurmanjî of 31, 129
 Turkî [Turki] (Iran-Turkic) 354, 398–401, 405–407, 412, 439, 441
 Turkic languages 3, 5, 7, 8, 11–16, 18–21, 24–26, 28, 29, 31, 33, 40, 157, 237–241, 260, 286, 337, 345, 354, 355, 357, 358, 360–371, 374, 376, 377, 382, 383, 385–390, 392–396, 398–401, 405–430, 432–439, 441, 443, 444, 446, 574, 575, 577, 578, 607, 610, 657, 821, 822, 825, 827–829, 854, 860, 931, 949
 Turkic, Iran 18, 24, 26, 360, 367, 371, 949

- Turkic, Iraq 14, 18, 383, 400, 409–411, 416, 418, 422, 423, 429, 434
 Turkish 3, 7, 8, 10, 13, 18, 20, 28, 40, 65, 72, 74, 77, 92, 95, 109, 121, 122, 138, 148–152, 156, 160, 162, 163, 176, 177, 181–184, 186, 188, 189, 195, 196, 228, 229, 235, 240, 259, 260, 264, 296, 303, 306, 309, 310, 315, 336, 345, 355, 357, 358, 360–366, 368, 371–374, 376, 377, 382, 383, 396, 399, 400, 406, 408, 413, 416, 418, 419, 421–425, 427, 431, 432, 436, 443, 444, 574, 575, 663, 822, 825, 827, 828, 830, 832, 833, 837, 839, 840, 855, 857, 892–895, 897, 899, 902, 908–911, 915, 917, 918, 920, 921, 923, 927, 929, 931, 942, 949
 Turkish, Ottoman (South Oghuz Turkic) 259, 345, 355, 363, 396, 400, 574, 895
 Turkish of Turkey (West Oghuz) 357, 399, 400, 406, 418, 422, 423, 425, 436
 Turkmanja [Turkman] 237, 354, 355, 357–359, 361, 363, 367, 374, 375, 377, 381, 383, 384, 398, 400, see also Turkish, Iraq
 Turkmen [Türkmen] 108, 354, 368, 382, 383, 398, 399, 405, 407
 Tûrûyô [Turoyo] 234–236, 305, 312
 Tutkâbonî 673–675, 700, 717, 728, 730, 731, 734, 762, 771, 824
 Txuma cluster of Christian NENA 191, 198, 218
 Tən, Christian NENA 232, 323
- U**
- Umra d-Shish, Christian NENA 224, 234, 324
 Urei 666, 668, 824
 Urmi, Christian NENA 17, 37–39, 191, 200–207, 210–215, 218, 222, 224–227, 229, 230, 494
 Urmi, Jewish NENA 191, 199, 202, 206, 210, 216, 218, 222–225, 227, 228, 486, 487, 526
 Uzbek 108, 570, 577, 607, 612, 616, 617
- V**
- Vafṣî [Vafsi] 25, 29, 468, 480, 577, 658–660, 665, 672, 676–678, 680, 681, 685–687, 691, 694, 695, 698–703, 705, 709–711, 713–715, 717, 719–723, 725, 728, 730–732, 738, 740, 742, 743, 746, 751–755, 757–760, 763, 765–770, 772, 773, 775–778, 780, 784, 789, 793–795, 797–805, 807–809, 815, 821, 822, 824, 948
 Vâjârgâhi 670, 743, 766
 Velâtrui 660, 666, 669, 670, 678
 Vidari 717
 Viznei 671, 690, 694, 712, 790, 824

W

Wārmāwa, Sorani of 278, 286

X

Xatirābādi [Khatirābādi, Xatirabadi] 666, 824

Xoşābari 664

Xərpa, Christian NENA 313, 323, 337

Z

Zakho, Jewish NENA 306, 313–315, 323, 324, 350

Zakho, Kurmanjî of 137, 287, 289–291

Zardayāna (variety of Gorani) 296, 298, 533, 540, 541

Zazaki 3, 106, 109, 149, 157, 158, 280, 281, 622, 623, 763, 837, 840

Zincirli, Kurmanjî of 151

Index of place names

A

- Adana (Turkey) 133
Adhaim (Iraq) 245, 297, 308
Adiyaman-Urfa (Turkey) 126
Afghanistan 354, 399, 444, 569, 570, 577, 579, 582, 623, 657
Ağrı (Turkey) 133
Ahmađ Āwā (Iraq) 534
Ahwaz [Ahvāz, Ahwāz, Ahwāz] (Iran) 305, 385, 481, 482, 484, 492, 504, 513, 514, 516, 525, 532, 949
Akbusak (Turkey) 133
Akre [Aqra, ḤAqra] (Iraq) 187, 264, 244, 288–291, 294, 351
al-Qamishli (Syria) → Qamishli
Alagir (Russia) 864
Alazani river 864
al-Ba‘āj [al-Ba‘āj] (Iraq) 245, 259
Aleppo (Syria) 3, 28, 184, 189, 260
al-Hasakah (Syria) → Hasakah
‘Alī Bāpīr (Iraq) 297
‘Alī Qūrchī (south of Arak) (Iran) 412
‘Alī Sarāy (Iraq) 297
Aligudarz (Iran) 445
Almaty (Kazakhstan) 195, 196
al-Mawṣil [al-Mauṣil] → Mosul (Iraq)
Alqosh (Iraq) 306, 312, 323, 324, 350, 352
Alvir (Iran) 711, 823
Amadiye [Amēdi, Amedia, Amediye] (Iraq) 16, 36, 37, 137, 290, 291, 306, 312–315, 317–319, 322, 486
Ambar 82, 297
Āmol (Iran) 661
Anatolia 3, 5, 8, 12, 13, 16, 25, 27, 28, 31, 34, 39, 41, 100, 106–109, 115, 123, 132, 134, 135, 150, 156, 159, 162, 170, 181, 182, 188, 190, 242, 244, 250, 303, 350, 355, 358, 361–363, 366, 367, 372, 383, 385, 386, 388, 398, 399, 401, 410, 413, 416, 421, 424, 428, 429, 436, 444, 825, 827, 828, 860
Anbarān (Iran) 663
Ankara (Turkey) 382, 444, 893
Ankawa (Iraq) 11, 239, 306, 313, 323, 324, 337, 350
Antep-Adiyaman (Turkey) 117
Anzali (Iran) 659
Aqra [ḤAqra] (Iraq) → Akre
Arab Emirates 569, 625, 657
Aradhin (Iraq) 306, 324

- Arbīl [Arbel, Arbil] (Iraq) 187, 237, 238, 240, 244, 264, 292, 297, 298, 306, 312, 314, 322–327, 335, 337, 339, 341, 343, 345, 351, 354, 355, 359–361, 370, 427, 947
Ardabil (Iran) 578, 664
Ardal (Iran) 447, 457, 460, 461, 467, 468, 471–473, 480, 947
Ardon (Russia) 864
Arhavi (Turkey) 832, 833, 835, 844, 852, 854, 860, 949
Armenia 31, 33, 65, 79, 80, 85, 87, 89, 93, 97, 98, 108, 123, 195, 196, 201
Arpačī (Iraq) 298
Artvin (Turkey) 831
Asālem (Iran) 660, 663, 823
Asia Minor 13, 27, 52, 573, 893, 903, 932
Asparēz 534, 536
As-Sulaymāniyya (Iraq) → Sulemaniyya
Astara (Azerbaijan) 660, 664
Australia 195, 310, 484, 569
Awbar (Iraq) → Ambar
Awroman [Awromān] (Iran) 296, 303, 566, 567
Azerbaijan 40, 190, 234, 355, 359, 362, 363, 365, 366, 383, 396, 398–400, 405, 406, 410, 412, 413, 416, 419, 424, 428, 429, 436, 481, 659, 660, 663, 664, 724, 791, 822
Āzex [Īdil] (Turkey) 189
- B**
- Ba’adra [Badr] (Iraq) 290, 294, 354
Ba’ashiqah (Iraq) 245
Badinan (Iraq) 116, 120
Baghdad [Bagdad] (Iraq) 194, 263–266, 354, 355
Bahzani (Iraq) 245
Bakragar (Iraq) 298
Baksan (Russia) 864
Bamō (Iraq) 298
Bāne (Iran) 297
Bānī Binok (Iraq) 534
Ba‘qūba (Iraq) 297
Bārāmāwā (Iran) 534
Bārānī Xwārū (Iran) 297
Bardī ‘Alī Xwārū (Iraq) 297
Barətla (Iraq) 306, 317
Barwar (Iraq) 16, 20, 45, 157, 238–240, 306, 312–315, 317–320, 322, 324–327, 330, 332–334, 337–340, 342–344, 346, 351, 500, 501, 949, 951
Barwari Bala (Iraq) 306

- Barxu (Iran) 297
 Barzan (Iraq) 306
 Başkale (Turkey) 191
 Batman (Turkey) 159
 Baxa Kon (Iraq) 534
 Bayâdestân (Iran) 19, 407, 409, 412, 413, 415, 416, 418, 419, 421–424, 427, 429–439, 441, 444
 Bayîj (Iraq) 297
 Bêdalan [Badalan] (Azerbaijan) 661
 Bêdyl (Iraq) 306, 310
 Beşköy (Turkey) 898–902, 914, 927, 930, 933
 Beslan (Russia) 864
 Betanure (Iraq) 306, 310, 312, 314, 323, 324, 336, 352
 Beth Arbaye (Turkey) 159
 Bêwyân (Iran) 297
 Bêwyânî Gawra (Iraq) 297
 Bičûl (Iraq) 297
 Bijar (Iran) 481
 Bîlbar 534
 Bingird (Iraq) 271, 285
 Bingöl (Turkey) 135
 Bisotun (Iran) 300
 Bitlis (Turkey) 97
 Bîyâra (Iran) 534
 Bîyâra (Iraq) 534
 Bohtan → Botan
 Bokan (Iran) 481, 496, 508
 Borb (Turkey) 90, 191
 Borîdar (Iran) 534
 Borûjerd (Iran) 297
 Botan [Bohtan] (Turkey) 14, 16, 32, 34, 35, 40–42, 107, 109, 158, 191, 194, 291
 Bukan [Bükân] (Iran) 297
 Bukhara (Uzbekistan) 569
- C**
 California (United States of America) 27, 97, 98, 102, 104, 195, 532
 Čamčamâl (Iraq) 298
 Čamšîdar (Iran) 534
 Caucasus 1, 3, 5, 8, 9, 26, 27, 85, 156, 195, 196, 568, 819, 820, 822, 825, 860, 862, 865, 884, 885, 890, 891
 Çaykara (Turkey) 893, 894, 897, 950
 Central Asia 72, 97, 100, 398, 573, 575, 622
 Central Province (Iran) 399, 406, 407
 Ceylanpınar (Turkey) 164
 Châl (Iran) 661
 Challa [Čâl] (Turkey) 191, 195, 233, 310
 Charaze (Iran) 661
- Chehr (Iran) 300
 Chehreqân [Chehreghân] (Iran) 661
 Chelgerd (Iran) 446
 Chenâr Rudxân [Chenâr Rudkhân] (Iran) 664
 Chikola (Russia) 862
 China 574, 601, 623, 698
 Cizre [Jazira] (Turkey) 107, 109, 187, 191, 269
 Crimea 52, 53, 92, 574
 Cudi mountain [Cudi Dağı] (Turkey) 191, 194, 195
- D**
 Dagâ Šêxân (Iraq) 534
 Dagâgâ (Iran) 534
 Dať 534
 Dałamarz 534
 Damascus (Syria) 260, 305
 Dandan Uiliq (China) 572
 Dâquq (Iraq) 297
 Dara (Iran) 534
 Dârâ (Iraq) 297
 Daragaz (Azeri exclave) (Iran) 399, 405
 Daragözü (Turkey) 35, 164, 170, 181, 187
 Dâraka (Iran) 297
 Darê Mař (Iraq) 534
 Darmûr (Iran) 534
 Darwêšân (Iran) 534
 Dastenâ (Iran) 461
 Dêr Molê (Iran) 534, 536
 Derigne (Iraq) 306
 Dersim [Tunceli] (Turkey) 53, 74, 94–96, 123, 127, 149, 156, 303
 Dêwaznâw (Iran) 534
 Dibba [Daba] (United Arab Emirates) 625, 627
 Dicle (Turkey) 107
 Digora (Russia) 862
 Dikin (-Marâqe) (Iran) 665, 687, 694, 698, 705, 729, 735, 773, 775, 779, 784, 822
 Diyana → Soran
 Diyarbakır [Diyarbakir] (Turkey) 106, 190, 421
 Diyarbakir Province 190
 Diz (Iran) 111, 945
 Dizâwar (Iran) 534
 Dizlî (Dezlî) (Iran) 534
 Dobe (Iraq) 306, 323, 337
 Dohuk [Duhok, Dohok] (Iraq) 11, 137, 138, 287, 289–291, 293, 303, 306, 314, 323, 947
 Dorud (Iran) 445
 Dukân (Iraq) 297, 298
 Dûšamnân (Iran) 297
 Dzagepbarz (Russia) 862

E

- Elāšt [Elāsh̄t] (Iran) 821
 Elaziğ (Turkey) 127
 Elbistan (Turkey) 121, 127, 136, 139, 141, 149
 Erbil → Arbil
 Erzincan (Turkey) 107
 Erzurum (Turkey) 127
 Esfahan → Isfahan
 Eslāmābād (Iran) 297
 Euphrates 1, 160, 237, 243, 248

F

- Fars, Province of 400, 416, 417
 Fereydun Shahr (Iran) 402
 Fuman (Iran) 659
 Furqān (Iraq) 298

G

- Gahvāre (Iran) 296
 Galēn (Iran) 534
 Galûgâh (Azeri exclave) (Iran) 399, 405
 Gandomāb (Iran) 672
 Gardabani (Georgia) 194
 Garja Kōyī (Iraq) 297
 Gawar [Gever] (official name Yüksəkova) (Turkey) 191
 Gawrajū [Gawrajū] (Iran) 296, 297, 302, 304, 534, 537, 567
 Gazakān (Iraq) 297
 Gāzārxān [Gozärkhon] (Iran) 674, 675, 820
 Gaziantep (Turkey) 151
 Georgia 9, 12, 65, 194, 195, 303, 388, 400, 413, 830, 861, 948, 949
 Germany 195, 553, 609, 618, 620, 893
 Geylavān (Iran) 662
 Gilan (Iran) 577, 578, 663–665, 821
 Girāt (Iran) 534
 Girdəni [Girdani] (Azerbaijan) 661
 Goṛgayī (Iran) 534
 Gōrī Ginūž (Iraq) 297
 Great Britain 573, 575
 Great Zab river 238, 306
 Gullī (Iraq) 291
 Gułp (Iraq) 534
 Gurān (5) (Iran) 661
 Gurchān (Iran) 665, 676

H

- Haftgel (Iran) 446
 Hājī Mīṣṭātā (Iraq) 297
 Hajj (Iran) 534

- Hakkari [Hakkâri, Hekari] (Turkey) 10, 108, 116, 158, 191, 193, 199, 201, 203–206, 209, 210, 212, 287
 Halabja [Halabja, Ḥalabja] (Iraq) 267, 295, 296, 301, 306, 351, 394, 397, 532, 534
 Hamadan (Iran) 406, 410, 416, 423, 578
 Hāna (Iraq) 534
 Hāna Garmâta (Iran) 534
 Harran (Turkey) 163, 164, 173
 Harsin (Iran) 27, 300, 302
 Harzan [Harzand] (Iran) 711
 Hasakah (Syria) 107, 161
 Hašamēz (Iran) 534
 Hashtpar (Iran) 660, 820
 Hasköy (Turkey) 164, 173, 183, 189
 Hawdiyan (Iraq) 314
 Hawrāmān [Hawraman, Hawromān, Awromān] (Iran) 296, 297, 385, 386, 533, 534, 536–538, 540, 541, 566, 568, 947
 Hawrāmān Žāwaro (Iran) 534
 Hawrāmān-i Luhon [Luhōn] (Iran) 296, 303, 534
 Hawrāmān-i Taxt (Iran) 296, 534
 Hayy 'Askarī (Iraq) 297
 Hezārrud (Iran) 672

I

- Ilam [Īlām] (Iran) 536
 India 95, 573–576, 621
 Iori river 864
 Iraf (district, Ossetia) (Russia) 862, 886, 887
 Iran 1, 8, 10, 12–14, 18–21, 24–27, 29, 31, 32, 34, 40, 65, 106, 108, 156, 190, 191, 195, 196, 198, 199, 202, 206, 213, 216–218, 222–224, 226–229, 238, 241, 267, 270, 273, 295, 296, 298, 302, 304–306, 309, 311, 316, 331, 332, 354, 358, 360–369, 371, 372, 377, 378, 383, 385–388, 392–396, 398–401, 405–407, 409, 410, 412, 413, 416–419, 421, 423–425, 427–430, 433, 436, 444–447, 478, 480–482, 484, 486, 507, 510, 516, 518, 525, 526, 532–534, 536–538, 566–580, 582, 608, 620–625, 627, 649, 657–660, 663, 664, 687, 693, 794, 796, 815, 819–824, 828, 932, 947, 949
 Iraq 1, 8, 12–14, 16, 18–21, 24, 25, 34, 36, 37, 41, 106–108, 116, 121, 126, 134, 137, 160, 162, 188, 190, 191, 194, 196, 199, 204–206, 209–215, 217, 218, 222, 223, 225, 237, 240–244, 246, 248, 252, 254, 263–265, 267, 269, 270, 273, 281, 287, 295, 296, 298, 299, 301, 305, 306, 309–317, 319–321, 323, 324, 329,

331–334, 336, 342, 345, 350, 352, 354, 355, 357, 358, 360–364, 366, 367, 369, 372, 377, 382, 383, 385, 388, 389, 395, 398–401, 407, 409–416, 418, 419, 421–429, 434–436, 444, 481, 482, 484, 487, 504, 510, 516, 518, 521, 533, 534, 537, 563, 569, 825, 827, 828, 837, 947, 949
 Isfahan [Esfahan] (Iran) 366, 385, 399–401, 406, 412, 415, 416, 423, 424, 438, 439, 445, 479, 762, 822
 Israel 195, 196, 246, 309, 311, 352, 484, 569
 Istanbul (Turkey) 157, 187, 195, 357, 383, 444, 860, 893, 910, 911, 915, 924
 Izeh (Iran) 445

J

Jamshidābād (Iran) 662
 Jazira → Cizre
 Jowkandān (Iran) 663
 Jub' Adin (Syria) 305
 Juban (Iran) 673

K

Kabarlū (Iraq) 297
 Kabate [Kafti] (Iran) 665
 Kabul (Afghanistan) 399, 412, 531, 607, 612
 Kajal (Iran) 711, 823
 Kalār (Iraq) 297, 298
 Kalāsur (Iran) 661, 672
 Kaljī 534
 Kamāla 534
 Kandelus (Iran) 669, 670, 820
 Kandūla [Kandüle] (Iran) 297, 534, 537
 Kānī Māz (Iraq) 297, 298
 Kānī Sīrna (Iraq) 297
 Karakoçan (Turkey) 122, 123
 Karan (Iran) 662
 Kařāwa (Iran) 534
 Kargānrud (Iran) 663
 Karnaq (Iran) 662
 Kars (Turkey) 120, 399
 Kaymna (Iran) 534
 Kayseri (Turkey) 133
 Kelārdašt [Kelardasht] (Iran) 659, 820
 Kelās [Kalās] (Iran) 584, 665
 Kerend (Iran) 481, 496, 497, 503, 508, 512, 532
 Keringān (Iran) 671, 823
 Kerman (Iran) 399
 Kermanshah [Kermanšāh, Kermānšāh, Kermānshāh] (Iran) 10, 27, 296, 302, 374, 385, 386, 394, 396, 416, 481, 534, 537, 541, 577, 578, 623

Kermanshah Province (Iran) 10, 27, 302, 385, 386, 394, 396, 481, 537

Khabor river 194
 Khanaqīn → Xanaqīn
 Khasab (Oman) 625, 627, 628, 947
 Khazhnidon (Russia) 862
 Khoīn [Xo'in] (Iran) 661

Khorasan [Khorāsān] (Iran) 108, 157, 400, 405, 582

Khorramshahr (Iran) 19, 21, 481, 482, 484, 487, 489–493, 504, 513–515, 517–522, 524, 525, 529, 531, 949, 951

Khoynarud (Azerbaijan) 661
 Khuzestān [Khuzestān] Province (Iran) 385, 445, 478, 479, 481, 484, 485, 490, 492, 947

Kifītī [Kifītī] (Iraq) 298, 421
 Kilit, Ordubad (Azerbaijan) 659, 664

Kinderib [Söyütlü] (Turkey) 188
 Kirkuk [Kirkük] (Iraq) 194, 265, 354, 355, 358–361, 380, 381, 383, 536

Kolur [Qolur] (Iran) 663
 Kordestan Province (Iran) 10, 385, 386, 394, 396, 406, 481
 Kordxeyl [Kordkheyel] (Iran) 668

Koy Sanjak [Koy Sanjaq, Koy Sanjaq] (Iraq) 296–298, 305, 306, 312, 313, 315, 316, 323, 352

Köya (Iraq) 297, 298
 Kozluk (Turkey) 13
 Krośnodar (Russia) 194
 Kūhdašt (Iran) 297
 Kuhrang (Iran) 457, 480
 Kumzar [Komzar, Kamzar] (Oman) 625, 627, 947
 Kur river 864

L

Lāhijān [Lāhijān, Lāhijān] (Iran) 659, 668, 820
 Lake Urmī [Urmī, Lake Urmia, Lake Urmīye] (Iran) 85, 190, 267, see also Urmī
 Lake Van (Turkey) 16, 108, 191
 Langerud (Iran) 659
 Larak Island (Iran) 625
 Laylān (Iraq) 298
 Lenkoran [Lənkəran] (Azerbaijan) 660, 663, 824
 Lerd (Iran) 662
 Lerik (Azerbaijan) 663, 716, 807, 814, 824
 Lesken (Russia) 862, 886, 887
 Levandavil (Iran) 663
 Little Zab river (Iraq) 245, 268
 Lordegān (Iran) 461
 Lotf-âbād (Azeri exclave) (Iran) 399, 405

M

- Ma’rib (Yemen) 625
 Mačiān [Māchiān] (Iran) 669, 670, 820
 Maçka (Turkey) 894
 Mahabad [Mahābād] (Iran) 191, 270
 Maka (Oman) 625
 Maku (Iran) 133
 Malā Ṣahmān (Iraq) 297
 Malatya (Turkey) 133
 Ma’lula (Syria) 305
 Mandali (Iraq) 14, 357, 363, 427
 Maraš [Mereš, Maraş, Maras] (Turkey) 52, 95, 96, 116, 148, 156, 166
 Mardin (Turkey) 14, 16, 31–35, 38, 39, 106, 116, 139, 141, 145, 152, 159, 160, 163–172, 174–180, 187, 189, 190, 196, 949
 Marīvān [Marīwān] (Iran) 534, 538, 541
 Māsāl [Masal] (Iran) 660, 664, 707, 711, 821, 951
 Masally (Azerbaijan) 659, 660, 663
 Masjed Soleymān [Masjed Soleyman] (Iran) 385, 445, 447, 449, 457, 947
 Māsule [Másule] (Iran) 660, 664, 820
 Ma’tiq (Iraq) 297
 Mazandaran (Iran) 577, 578, 820
 Māzībin (Iran) 534
 Mazun (Oman) 625
 Medina (Saudi Arabia) 625
 Mesopotamia 1, 4, 65, 159, 160, 237, 243, 269, 305, 352, 355, 571, 825, 827
 Mēxās (Iraq) 297
 Midin (Turkey) 208, 209, 949
 Midyat (Turkey) 107, 133, 144, 148, 152, 160, 171, 190, 195, 207–209, 213
 Mlahso [Mlahso, Mlaħsô] (Turkey) 32, 41, 42, 45, 190, 195, 197, 198, 204–207, 209, 212, 213, 220, 221, 234, 526
 Miğuryānī Ambar (Iran) 297
 Mosul (Iraq) 23, 222, 240, 243, 244, 246–250, 252–260, 269, 270, 295, 296–298, 306, 309–311, 325, 334, 340, 345, 351, 352, 355, 357, 533, 947, 949
 Muš [Muş] (Turkey) 8, 39, 50, 52, 54, 61, 62, 79, 80, 90, 94, 100, 102, 132, 159
 Musandam (waliyat, province) (Oman) 385, 625, 627, 658, 948, 951
 Musandam peninsula (Oman) 625
 Mutki (Turkey) 182

N

- Naltchik (Russia) 864
 Narwē (Iran) 534
 Nāvrud (Iran) 663

- Nāwa (Iran) 534
 Nawsūd (Iran) 385, 534, 536–539, 541, 542, 544, 548, 549, 554
 Naysāna (Iran) 534
 Nerwa (Iraq) 11, 306, 310, 312–314, 323, 324, 353

- Nineveh plains (Iraq) 296
 Nodšā [Notšā] (Iran) 534, 537, 541, 549
 Noubarān (Iran) 435, 441, 442
 Nowkiān (Iran) 662
 Nowṣūd (Iran) 297
 Nūrābād (Iran) 297
 Nuşaybīn [Nasibin] (Turkey) 159
 Nwēn 534

O

- Of (Turkey) 892, 894, 896
 Oman 385, 625, 636, 638, 649, 656–658, 947, 951
 Oskulak (Iran) 664, 665, 673, 786, 824

P

- Pakistan 569
 Para (Turkey) 8, 33, 128, 202, 314, 368, 379, 902, 916, 928
 Pareh Sar (Iran) 662
 Paveh [Paveh, Pāwā, Pāvē] (Iran) 296, 297, 534, 537, 568, 573
 Pāyrvand (Iran) 300
 Persian Gulf 623, 657, 717
 Persis [Fārs] (Iran) 571
 Peshtətuk [Peshtatuk] (Azerbaijan) 661
 Pīrānšahr (Iran) 297
 Pirasora [Pirəsorə] (Azerbaijan) 756, 814
 Piždar (Iraq) 271, 285

Q

- Qabul → Kabul
 Qādir Karam (Iraq) 398
 Qałāgā (Iran) 534
 Qałajē (Iran) 534
 Qałama (Iraq) 297
 Qal’at Ḥasan’ābād (Iran) 482
 Qaleh Gözäldereh [Qal’ä Gözälđärä] (Iran) 424
 Qāmīš (Iraq) 297, 298
 Qamishli (Syria) 133, 268, 307
 Qarāmīn (Iraq) 297
 Qaraqosh (Iraq) 10, 11, 240, 306, 309, 311–313, 315–324, 326, 328, 329, 333, 334, 337, 340–345, 351
 Qarluq (Iran) 402

- Qarṭmīn (Yavantepe) (Turkey) 36, 37, 43, 44, 173, 180, 181
 Qaṣqa (Iraq) 298
 Qasr-e Ṣirin [Qaṣr-e Ṣirīn] (Iran) 481
 Qazvīn [Qazvin, Qazvīn] (Iran) 399, 400, 410, 482, 485, 578, 664, 665, 673
 Qolur → Kolur
 Qom (Iran) 406
 Qorve [Qorveh] (Iran) 945–947, 949
 Qorveh (Iran) 945–947, 949
 Quetta (Pakistan) 569
 Qūṭe (Iraq) 534
- R**
 Ḥamāzān (Iraq) 297
 Rāmsar (Iran) 659, 668
 Rasht (Iran) 659
 Razajerd (Iran) 661
 Ḥazāw (Iran) 534
 Rewandiz [Ruwanduz] (Iraq) 292, 306, 323
 Rezvānshahr (Iran) 662
 Rostamābād (Iran) 662
 Rudbār [Rudbār] (Iran) 659, 660, 664, 665, 667, 673, 820
 Ruma (Turkey) 14, 34, 42, 191, 203, 207, 216–218, 228
 Rustaqā (Iraq) 306, 323, 335–337, 351
- S**
 Sablagh → Mahabad
 Sagzābād (Iran) 661
 Sainqala (Iran) 481
 Salamas [Salmās] (Iran) 13, 34, 40, 191, 195, 204, 235, 396
 Samarkand (Uzbekistan) 569
 Šammās (Turkey) 191, 202–204, 213
 Šāmyān (Iran) 534
 Sanandaj [Sanandaj] (Iran) 10, 15, 19–21, 23, 269, 306, 351, 386, 387, 389–394, 397, 481, 482, 484–497, 499–501, 503–512, 516–519, 521–524, 526, 527, 532, 534, 577, 949, 951
 Sandu (Iraq) 310
 Ṣanlūrfa [Urfā, Edessa] (Turkey) 159, 162–165, 173, 182, 184, 188, 264, 416, 949
 Sāqeqz [Saqqeqz] (Iran) 481, 494, 497
 Sarachos (Turkey) 897, 898, 900, 902–904, 920
 Sar Āqā Seyyed (Iran) 448, 449
 Sar Pol-e Zahāb (Iran) 296, 297, 533
 Sarak [Sərək] (Azerbaijan) 664
 Sardasht (Iran) 107
 Sargat (Iraq) 534
 Sāri (Iran) 821, 823
 Sār-i Hawrāmān (Iran) 534
 Sarqaṭā (Iran) 297
 Sarqizil (Iraq) 297
 Sar-u Pīrī 534
 Ṣarya (Iraq) 290
 Sason (Turkey) 13
 Ṣawak (Iraq) 298
 Say Miṣafā (Iraq) 297
 Ṣāy Tōtyā (Iran) 297
 Sayyādlar [Sayyadlar] (Iran) 663, 712
 Sē Girdkān (Iraq) 297, 298
 Ŝemzinān [Şemdinli, Şemzinan] (Turkey) 13, 33–35, 41, 108, 120–122, 125, 132, 135, 137–139, 144–146, 148
 Senna (Iran) → Sanandaj
 Ŝəno (official name Ushnuye) (Iran) 191, 213
 Sərək [Sarak] (Azerbaijan) 661
 Ŝexān (Iran) 534
 Shahe (Iraq) 310
 Shahrān (Iran) 662
 Shahr-e Kord (Iran) 445, 479
 Shāl (Iran) 662
 Shāndermen (Iran) 664
 Shapat (Turkey) 314
 Shaqlawa (Iraq) 306, 323, 334
 Sheker (Russia) 862, 886, 887
 Shino (Iran) 107
 Shiraz [Širāz] (Iran) 400, 434
 Shirqat (Iraq) 245
 Shwata (Turkey) 191
 Siāvarud (Iran) 662
 Siirt (Turkey) 13, 35, 159, 162, 165, 170–172, 187, 191, 234
 Silēn (Iran) 534
 Sinjar (Iraq) 244, 290
 Sinjar mountains (Iraq) 244
 Sipasar (Iraq) 298
 Širnak (Turkey) 187, 191
 Sivas (Turkey) 106
 Siverek (Turkey) 173, 188
 Solduz (official name Naghade) (Iran) 191
 Sonqor (Province of Kordestan) (Iran) 368, 399, 406, 407, 409, 412, 413, 415, 416, 421, 422, 428, 429, 434, 437–441, 949
 Soran [Diyana] (Iraq) 133
 Sosakān (Iraq) 534
 Šošmē (Iran) 534
 Soviet Union 195, 570
 Sufaya (Iraq) 297
 Sulak river 864

Sulemaniyya (Iraq) 15, 107, 270–273, 277–279, 297, 305, 306, 312, 313, 315, 319, 321, 323, 324, 327, 330–335, 339–342, 344, 345, 351, 394, 397, 481, 482, 487, 505, 510, 532, 535
 Surčī [Surčī] (Iraq) 281, 291
 Sürmene (Turkey) 892–895
 Sweden 195, 196, 484

T
 Tabaristan (Iran) 572, 622
 Tabriz (Iran) 12, 388, 400, 406, 407, 413, 416, 820
 Tahārom (Iran) 662
 Tajikistan 569, 570, 577, 579, 582
 Tākestān (Iran) 663, 664
 Tal’āfer [Talafer, Telafer, Tel’āfer] (Iraq) 354, 355, 358
 Tālāwān (Iran) 534
 Tāleš-Dolab (Iran) 664
 Tallkayf [Tall Kayf] (Iraq) 246
 Tangī Hamām (Iran) 297
 Tapa Čarmē (Iraq) 297
 Taqtaq (Iran) 297, 298
 Tāw Doław (Iran) 534
 Tawēla (Iraq) 534
 Tawrēbar (Iran) 534
 Tāzāwā (Iran) 534
 Tbilisi (Georgia) 45, 195, 303, 860, 890
 Tehran (Iran) 195, 401, 431, 478–480, 484, 541, 566, 567, 578, 581, 596, 613, 615, 621, 624, 663, 704, 819–823
 Telkepe (Iraq) 306, 318, 323, 324, 334, 350, 352
 Terek River 864
 Tergawar (Iran) 191
 Tēžtēž, (Iran) 534
 Tifliya (Iran) 534
 Tigris 1, 10, 19, 23, 31, 103, 190, 194, 228, 237, 243, 244, 246, 247, 252, 305, 313, 827
 Tikrit [Tikrīt] (Iraq) 264, 309
 Tillo (Turkey) 160, 162–164, 177, 178, 180–183
 Toldzgun (Russia) 862
 Tonekabon (Shahsavār) (Iran) 659
 Tonya (Turkey) 893, 894
 Topzawa [Tōpzāwa] (Iraq) 298, 301
 Trabzon (Turkey) 72, 893, 897, 906, 931, 933, 950
 Tskhinval (South Ossetia, Georgia) 864
 Tūlaban (Iraq) 297
 Tulaguvan [Tulaguvan] (Azerbaijan) 661
 Tunceli → Dersim

Tur Abdin [Tūr ‘Abdīn, Tur ‘Abdīn] (Turkey) 129, 162, 181
 Turkey 1, 10, 14, 31, 40, 65, 72, 95, 102, 104, 106–109, 114, 116, 122, 137, 145, 156, 157, 159, 162, 187, 188, 190, 191, 193–196, 198, 204, 218, 228, 264, 287, 290, 291, 303, 305, 306, 309, 310, 312, 314, 315, 355, 357, 383, 398–400, 406, 418, 422, 423, 425, 436, 481, 482, 569, 570, 825, 830, 832, 860, 862, 892, 893, 931, 932, 949, 950
 Tutkābon (Iran) 662
 Tūz Xurmātū [Tuzkhurmati, Tuzkhurmati, Tuz Khurmatu] (Iraq) 355, 358–360, 376, 378, 421
 Tuzlagözü (Turkey) 170

U
 United Arab Emirates 625, 657
 Ure (Iran) 660, 669, 670
 Urfa → Şanlıurfa
 Urmi [Urmia] (Iran) 9, 12, 13, 17, 18, 28, 33, 34, 37–40, 42, 45, 59, 64, 81, 82, 85, 86, 94, 190, 191, 194–196, 198–207, 210–216, 218, 222–230, 235, 267, 309, 311, 316, 351, 352, 386–388, 395–397, 413, 481, 482, 486, 487, 494, 526, 949, 951, see also Lake Urmi
 Ursdon-Sindziqaw (Russia) 863
 USA (United States of America) 195, 484, 569
 Usarā’ al-Mafqūdīn (Iraq) 297
 Uzbekistan 569, 570
 Uzungöl → Sarachos

V
 Vafs (Iran) 663, 676, 699, 703, 759, 768, 769, 794, 809
 Vājārgāh (Iran) 670
 Van (Turkey) 1, 7, 15, 16, 22, 24, 33, 51, 52, 59, 64, 80–85, 90, 93, 97, 98, 102–104, 106–108, 117, 191, 387, 389–391, 395, 421, 457, 458, 480, 625, 628, 632, 639, 649, 656, 891, 948
 Varto (Turkey) 133
 Varzob (Tajikistan) 607
 Velatru [Velātru] (Azerbaijan) 821
 Vidar (Iran) 823
 Viranşehir (Turkey) 164
 Vizne (Iran) 663
 Vladikavkaz (Russia) 863, 889, 890

W
 Wallagkom valley (Russia) 862
 Wardak (Iraq) 297
 Wārmāwa (Iraq) 278, 286
 Wasē (Iran) 534

Waysīā 534

Wēlā (Iraq) 298

West Azerbaijan province (Iran) 190, 396, 481

X

Xānagā (Iran) 534

Xanaqīn [Khanaqin, Xānaqīn] (Iraq) 295–298,
306, 357, 427

Xatirābād [Khatirābād] (Iran) 661

Xirosa (Iran) 534

Xorramābād (Iran) 297

Xošābar (Iran) 664

Xošāw (Iran) 534

Xurmāl 534

Y

Yemen 625

Yengi Tis‘īn (neighbourhood in Kirkuk)
(Iraq) 358, 360

Yerevan (Armenia) 45, 92–105

Yünlüce → Mlahso

Z

Zab → Great Zab river

Zagros (mountains) (Iran) 269, 385, 445, 478,
533, 947

Zakho (Iraq) 137, 287, 289–291, 306, 313–315,
323, 324, 350

Zalka (Iran) 534

Załm (Iraq) 534

Zanjan [Zanjān] (Iran) 410, 578

Zaqar (Iraq) 297

Zarda [Zarde] (Iran) 296, 297, 304, 534, 537, 567

Zébārī (tribal confederation) 288

Zincirli (Turkey) 151

Žīwār (Iran) 534

Zom (Iran) 534

Žūrū (Iraq) 297

Index of subjects

A

- ablative 68, 83, 364, 372, 382, 415, 418, 419, 429, 430, 443, 594, 707, 807, 818, 835, 838, 868, 876, 888, 919
accusative 44, 68, 77, 79, 83, 92, 131, 207, 217, 221, 241, 330, 364, 382, 393, 395, 408, 418, 419, 426, 443, 526, 772, 778, 780–783, 818, 826, 827, 852, 868, 888, 902–906, 919, 921, 931
accusative (alignment) 44, 330, 331, 334, 393, 395, 526, 770, 771, 784
accusative (case) 780, 781, 783, 919
active (voice) 42, 169, 171, 186, 188, 219, 220, 251, 255, 263, 328, 332, 393, 506, 513, 632, 927
adjective 37, 38, 118, 173, 220, 369, 389, 419, 428, 429, 465, 466, 477, 504, 514, 517, 519, 543, 545, 548, 565, 611, 612, 630, 631, 643, 701, 702, 769, 770, 839, 902, 904, 918
adjective, comparative → comparative
adjective, superlative → superlative
adposition 115, 120, 131, 155, 282, 283, 394, 428, 544, 696, 705, 760, 764, 805, 818, 826
adverb 419, 456, 545, 548, 551, 585, 593, 607, 630, 631, 644, 685
adverbial 226, 227, 343, 371, 395, 396, 429, 454, 470, 471, 525, 586, 596, 631, 647–649, 687, 795, 796, 798–802, 820, 851, 856, 857, 897, 917, 928
adverbial clause 226, 227, 371, 395, 396, 470, 471, 525, 647–649, 687, 796, 798–802, 820, 856, 857, 928
agent 44, 241, 372, 373, 394, 420, 422, 423, 429, 432, 544, 547–550, 552, 609, 631, 689, 696, 727, 773–779, 782, 786
agentive 218, 332, 505, 507, 508, 513
agreement 44, 116, 125, 126, 131, 145, 221, 241, 261, 264, 267, 273, 282–286, 291, 294, 298, 315, 325, 331, 510, 520, 543, 565, 584, 603, 776, 777, 779, 822, 826, 841–843, 847, 848, 852, 872, 875, 876, 901, 904, 911, 918, 926, 932
agreement, in number 116, 291, 294, 584, 603
agreement trigger 835, 842, 843, 846–848, 852, 872
alignment 6, 28, 44, 131, 156, 241, 285, 303, 330–332, 334, 335, 345, 350, 383, 393–395, 464, 478, 507, 525, 526, 552, 553, 689, 693, 698, 711, 756, 770–782, 784, 820, 826, 837, 852, 876, see also accusative (alignment), double oblique alignment
alignment splits 770, 772
allophone 9, 197, 312, 360, 362, 408, 409, 411, 448, 449, 485, 539, 581, 582, 677–679, 832
analytical attribution 253
animacy 585, 597, 697, 708, 757, 770, 775, 867, 901, 904, 932
aorist 70, 71, 74–76, 78, 80, 89, 91, 365–367, 369, 372, 382, 396, 421–424, 430, 443, 840, 845, 848, 859, 888, 911–915, 920, 926, 931
applicative 847–849, 853, 859, 873
Arabic script 270, 417, 570, 572, 573, 607
areal convergence 9, see also convergence, contact
areal diffusion 8, 10, 12, 13, 15, 27, 29, 156, 502, 526, 822, 829, 860
aspect 41, 126, 138, 140, 145, 149, 183, 220, 221, 252, 282, 302, 371, 372, 391, 396, 420, 431, 438, 439, 456, 457, 513, 514, 532, 549, 632, 633, 636, 637, 642, 648, 657, 733, 827, 840, 871, 891, 912–916, see also imperfective, perfective, progressive
aspiration 9, 316, 363, 410, 412, 540
assimilation 13, 83, 179, 210, 238, 271, 300, 363, 364, 407, 408, 414–416, 418, 420, 421, 494, 495, 509, 543, 579, 630, 719, 866, 894
attributive 27, 37, 43, 44, 239, 372, 389, 397, 433, 504, 545, 851
auxiliary 129, 149, 170, 332, 341, 391, 422, 437, 460, 461, 477, 519, 603, 609, 636, 738, 767, 818, 873, 915, 930, 931

B

- benefactive 18, 122, 123, 143, 282, 284, 369, 393, 594, 597, 636, 805, 807, 843, 848, 849, 916
bilabial(s) 113, 137, 164, 247, 287, 361, 407, 408, 409, 410, 414, 540, 627, 629, 865
bilingualism 181, 196, 246, 300, 577, 578, 677, 704
binding (syntax) 151, 878
borrowing 6–8, 18, 26, 28, 29, 150, 151, 156, 157, 240, 251, 345, 351, 532, 566, 583, 664, 689, 865, 884, 885, 895
breaking of vowel 319, 388, 491, 492
broken (Arabic) plural 7, 8, 586, 587

C

case 15–17, 37, 40, 44, 50, 73, 114–116, 118, 119, 120, 131–133, 135, 137, 147, 166, 239, 249, 273, 285, 291, 293, 301, 331, 332, 364, 365, 368, 370–373, 377, 394, 396, 413, 417–419, 426–429, 431, 432, 437, 451, 553, 543–545, 547, 553, 571, 585, 587, 594, 631, 687, 689–691, 694, 696–701, 709, 710, 755, 756, 763, 764, 770, 772, 774, 778–783, 791, 793, 794–796, 808, 825, 826, 827, 834–838, 841–843, 846, 847, 850–854, 866–868, 870, 875, 876, 877, 885, 900, 902, 904–906, 919, 929
 case, see also ablative, accusative, dative, genitive, instrumental
 case form 37, 691, 700, 709, 778–781
 case marking 114, 115, 116, 166, 239, 249, 331, 332, 370, 394, 417, 451, 533, 689, 698, 772, 783, 793, 795, 825–827, 847, 852, 854, 866, 868, 870, 875, 929
 case morphology 16, 17, 364, 377, 418
 case suffixes 147, 368, 372, 373, 415, 417, 426, 429, 431, 432, 838, 868
 case system 585, 690, 691, 694, 696, 819, 827, 868, 885
 causative 124, 169, 330, 351, 505, 513, 548, 690, 765, 768, 769, 818, 840, 846, 849, 859, 873, 887, 917, 921
 Christians 12, 28, 45, 99, 160, 162, 163, 190, 191, 194, 195, 235, 244, 246, 305, 309–311, 314, 344, 352, 357, 397, 481, 482, 485, 894, 920, 951
 citations 468, 471
 clause chains 636, 647–649
 cleft construction 337
 clitic 15, 17–19, 22, 35–37, 44, 117, 118, 125, 126, 130, 135, 147, 150, 240, 251, 273, 274, 279–283, 285, 286, 293, 299–302, 304, 330, 368, 369, 391, 426, 434, 452, 464, 477, 512, 515, 526, 541, 546, 547, 552–554, 556, 584, 591–594, 603, 730, 731, 786, 789, 790, 794, 827, 854, 872, 877, 888, 905–908, 911, 923, 924, 931, 932
 clitic doubling 877, 932
 clitic pronoun → pronoun, clitic
 coda 319, 449, 450, 629, 630, 634, 652, 653, 899
 code-switching 12, 151, 156, 311, 345, 621, 832, 886, 887
 comparative (adjective) 7, 8, 151, 370, 419, 426, 584, 585, 589, 632, 643, 923, 929
 complement 18, 22, 115, 120, 122, 123, 125, 226, 227, 276, 281, 283, 284, 337, 340, 344,

371, 372, 375, 394, 429, 435, 436, 455, 471, 544, 547, 548, 551, 591, 594, 601, 614, 636–639, 641, 642, 647, 648, 650, 795, 796, 802, 803, 806, 843, 853, 855–857, 875, 881, 919, 922–924, 926–928
 complement clause 226, 227, 276, 284, 344, 371, 372, 375, 429, 435, 471, 647, 648, 650, 803, 855, 856, 881, 922–924, 928
 complementizer 155, 284, 302, 471, 477, 553, 621, 829, 855, 859, 881, 888, see also complement clause
 complex predicate 22, 130, 134, 151, 156, 171, 273, 293, 912, see also light verb
 compound 275, 302, 377, 437, 450, 454, 514, 520, 545, 546, 565, 602–604, 608, 609, 624, 632, 635, 636, 638, 645, 702, 706, 734, 737, 738, 755, 768, 769, 839, 908
 compounding 702
 compound marker 546, 565
 conditional 70, 149, 150, 155, 177, 187, 227, 258, 279, 292, 302, 344, 345, 382, 435, 438, 439, 444, 523, 525, 552, 608, 610, 618–620, 649, 688, 723, 724, 729, 732, 742–746, 748–752, 796, 797, 800, 801, 818, 837, 857, 871, 897, 930, 931
 contact, language → language contact
 contact-induced change 6, 22, 25, 26, 31, 150, 358, 361, 363, 368, 385, 386, 401, 407, 410–412, 479, 860, 903, 923
 contact-phenomena 181, 259
 continuous 6, 51, 69, 96, 171, 252, 293, 350, 457, 551, 578
 converb 871, 873, 888
 convergence 6, 9, 13–15, 17, 25, 29, 35–38, 40, 42, 44, 45, 228, 238, 240–242, 298, 316, 321, 332, 345, 383, 385, 388, 390, 393–396, 407, 444, 446, 447, 490, 526, 534, 687, 821, 822, 829, 891
 coordination 465, 468, 469, 620, 646, 647, 657, 658, 821, 879
 copied Iranian specifiers in {+AkA} or {+i} 15, 368, 374, 389, 426, 434, 543, 545
 copied Iranian enclitic personal pronouns 368, 426
 copula 18, 20, 39–44, 50, 51, 69, 74, 78, 86, 87, 92, 126, 139, 140, 145, 151, 155, 172, 181, 183, 186, 207–211, 216–218, 220–223, 235, 240, 252, 253, 263, 266, 301, 302, 320, 333–339, 341, 351, 366, 382, 390–392, 414, 416, 420–422, 424, 425, 443, 455, 460–462, 509–512, 515–519, 526, 547, 551, 552, 565,

- 584, 603, 605–607, 613, 614, 621, 695, 718, 725, 733, 734, 757, 758, 783, 790, 803, 818, 827, 873, 888
 copying 8, 182, 368, 433, see also borrowing, copied
 counterfactuals 743, 751, 919, 920, 926
- D**
- dative (case) 18, 68, 83, 92, 330, 332, 364, 367, 368, 370–372, 382, 393, 394, 418, 419, 426–428, 431, 443, 505, 515, 795, 826, 835, 837, 838, 842, 843, 847–849, 852, 853, 859, 868, 873, 876, 888, 933, 948
 definite article 17, 35, 92, 172–174, 179, 180, 182, 186, 207, 209, 212, 213, 263, 274, 322, 323, 368, 389, 426, 452, 501, 502, 588, 900, 903, 904, 906–908, 918, 919, 931
 definiteness 15, 16, 21, 114, 239, 273–275, 284, 389, 451, 452, 465, 478, 543, 545, 546, 585, 588, 642, 703, 775, 903, 932
 deictic copula 39, 41, 43, 209, 211, 334
 deixis 16, 35, 39, 208, 210, 211, 321, 322, 388, 389, 496–498
 delabialization (of rounded vowels) 363, 396, 411
 demonstrative 16, 17, 35–39, 116–118, 120, 122, 139, 141, 155, 184, 186, 214, 234, 263, 274, 275, 284, 300, 302, 304, 323, 324, 327, 364, 365, 374, 375, 382, 388, 389, 394, 408, 419, 443, 452, 454, 455, 462, 465, 496–498, 503, 542, 543, 545–547, 554, 555, 565, 588, 590, 592, 612, 616, 621, 630, 631, 643, 656, 692, 693, 733, 818, 837, 838, 852, 859, 882, 883, 905–907, 918, 931
 demonstrative pronoun 16, 17, 35, 38, 214, 234, 300, 323, 324, 327, 364, 365, 374, 375, 388, 389, 394, 408, 419, 455, 462, 496, 497, 503, 546, 547, 5590, 592, 631, 837, 838, 905, 906
 derivational verbal stem 215, 219, 220, 328, 329
 determiner 116, 284, 902
 deverbal 632, 633, 635, 638, 642, 643, 645, 648
 devoicing 61–64, 90, 112, 127, 165, 248, 492, 540, 543, 866, 900
 devoicing of consonants 248
 dialect 8–10, 12, 13, 18, 19, 21, 23, 25, 28, 32–35, 40, 42, 44, 45, 47, 48, 51–53, 62–66, 69, 70, 72, 73, 75, 77, 79–81, 85–89, 92–105, 107–109, 116, 121, 122, 125, 129, 132–134, 137–139, 141, 145–147, 149, 152, 156, 157, 159, 160, 163, 164, 169, 170, 172, 173, 177, 179, 181–183, 187, 188, 190, 191, 195–201, 203, 210, 217, 220, 224, 230, 233–235, 241, 243, 244, 247–249, 252–254, 259, 260, 263–266, 270–273, 281, 285–290, 294, 296, 303, 306, 310, 311, 313–315, 317, 320, 322–324, 328, 335, 339, 343, 345, 346, 350–352, 361, 368, 382, 383, 385, 387, 390–392, 394, 396–398, 412, 413, 444, 445, 447–449, 457, 467, 468, 471–473, 479–482, 487, 492, 493, 504, 505, 510, 514–516, 521, 524–527, 529, 531, 532, 534, 567, 577, 607, 628, 656–659, 663–667, 670, 673, 676, 711, 724, 757, 820–824, 832, 860, 862, 863, 878, 889, 890, 897, 932, 933, 949, 951
 dialect continuum 107, 134, 673
 diaspora 106, 109, 157, 234, 237, 246, 385, 574, 892, 893
 differential object marking (DOM) 302, 462, 689, 6693, 770–773, 775–777, 867, 876
 diphthong 13, 34, 62–64, 204, 210, 318, 491, 492, 494, 495, 497, 581
 direct case 115, 116, 120, 131–133, 147, 293, 544, 696, 755, 774, 780, 781, 818
 direct object 5, 18, 21–26, 130–132, 169, 176, 221, 223, 240, 241, 250, 254, 280, 281, 283, 301, 302, 333, 334, 345, 368, 373, 394, 426, 435, 452, 464, 465, 467, 518–520, 547, 548, 588, 590, 591, 595, 597–600, 614, 635, 638, 693, 710, 827, 854, 868, 876, 877, 903, 906, 921
 direct object marking 467, 590, 595, 597
 ditransitive 791, 793, 921
 documentation 295, 658, 832, 860, 892, 894, 933
 double oblique alignment 689, 773, 775–780, 782, 789
 drawl (suprasegmental intonation pattern) 367, 426, 433
 dual number 166, 174, 186, 249, 263, 588, 632
 durative 252, 584, 606, 611, 621, 712, 715–718, 729–731, 734, 736, 818

E

- echoic 839, 840, 859
 ejective 10, 825, 865, 884, 885
 emphatic 315–317, 367, 408, 419, 452, 583, 593, 605, 621, 627, 628, 631, 646, 818, 888
 enclitic indefinite marker 182
 endangered language 3, 163, 246, 270, 310, 385, 484, 533, 822, 825, 832, 892, 893, 933
 equative (case) 338, 365, 372, 418, 419, 868, 888
 ergative 6, 44, 131, 132, 156, 241, 294, 331, 332, 345, 393–395, 507, 508, 510, 511, 526, 689, 711, 756, 773–780, 826, 835, 837, 838, 852, 853, 859, 885

- ergative agreement 241
 ergative alignment 44, 331, 332, 345, 393,
 394, 526, 711, 773, 774
 ergativity 6, 108, 131, 132, 156, 157, 235, 241,
 285, 332, 350, 351, 394, 395, 531, 532, 689,
 711, 773, 774, 778, 827, 828, 891
 evidentiality 183, 395, 479, 639, 656, 840, 844,
 871
 evidential/inferential copula in {-ImI\\$} 366,
 424
 experiencer 150, 220, 286, 287, 547, 551, 603,
 693, 696, 755, 756, 790, 843, 876
 experiencer verbs 755, 790
 ezâfe [ezafe, ezâfe, izafe] 17, 37, 38, 40, 42–44,
 114, 115, 117–121, 133, 137, 139–142, 145–
 147, 149, 150, 152, 155, 156, 214, 239, 240,
 273, 275, 284, 291, 293, 294, 302, 303, 327,
 371, 382, 394, 428, 429, 443, 451, 453, 454,
 465, 466, 477, 503, 504, 523, 531, 533, 544–
 546, 565, 572, 584, 589, 594–596, 601, 602,
 611, 612, 617, 621, 624, 689, 691, 697–702,
 806, 818, 819
 ezâfe construction 44, 115, 117–120, 133, 139,
 239, 275, 371, 418, 428, 504, 697, 699, 701
- F**
- feminine (gender) 16, 114, 115, 119, 141, 155,
 167, 172, 174, 179, 180, 186, 213, 214, 248,
 249, 253, 263, 321, 324, 325, 500–502, 544,
 545, 548, 551, 552, 554, 555, 557, 565, 624,
 678, 684, 685, 689, 702, 818, 819, 899–904,
 918, 931
- finite 18, 50, 125, 140, 168, 171, 241, 250, 251,
 255, 277, 328, 369, 371, 374–376, 396, 431–
 433, 435, 438, 439, 456, 457, 459, 548, 549,
 551, 552, 607, 615, 617, 627, 635, 685, 725,
 827, 828, 833, 835–837, 840, 843, 845, 850,
 851, 854, 856, 871–873, 876, 881, 911, 915,
 917, 922–924
- focal present 365, 366, 396, 421–423
- focus 3, 4, 20, 151, 164, 186, 209, 219, 237, 243,
 338, 428, 446, 457, 516, 533, 600, 663, 687,
 759, 839, 889, 897, 921
- fronting (of vowels) 13, 15, 34, 61, 81, 90, 137,
 204, 238, 290, 318, 410, 413, 614, 756, 764,
 777, 786–791, 921
- fronting (of stops and affricates) 271, 359, 360,
 407
- future 4, 26, 41, 50, 99, 109, 125, 126, 128, 129,
 147, 149, 155, 170, 183, 186, 218, 221, 223,
 224, 239, 252, 276, 278, 292, 295, 302, 332,
 333, 340, 367, 423, 425, 435, 447, 511, 515,
- 518, 570, 577, 598, 603, 605, 606, 609–611,
 615–619, 621, 628, 633, 665, 678, 688, 712,
 713, 715, 716, 718, 724–727, 737, 781, 790,
 818, 845, 850, 851, 859, 871, 873–875, 881,
 888, 911–914, 920, 925, 930, 931
- G**
- geminate 583
- gender 15, 16, 35, 108, 114, 117–120, 139–142,
 146, 147, 157, 167, 180, 207, 208, 210, 239,
 249, 254, 273, 291, 293, 301, 321–326, 330,
 335, 337, 389, 451, 497, 501, 502, 510, 515,
 516, 533, 543–545, 550, 552, 555, 557, 571,
 585, 631, 685, 687, 689, 690, 692, 694, 699,
 701, 773, 777, 822, 825, 835, 867, 900–902,
 904–907, 909, 918, 932
 gender, grammatical 114, 273, 451, 533, 543,
 544, 689, 692, 867, 900, 901, 904
 gender, biological 114
- genitive (case) 17, 36–38, 44, 51, 68, 73, 74,
 83, 92, 186, 187, 189, 212, 214, 239, 253,
 263–265, 327, 364, 373, 382, 394, 416–419,
 426, 443, 500, 503, 504, 524, 544, 689, 691,
 694–696, 699, 710, 755, 764, 779, 818, 835,
 838, 850, 851, 859, 868, 902, 903, 908, 931
- glottalization 8–10, 825, 826, 833
- goal 5, 23–26, 115, 122, 130, 131, 241, 280, 283,
 340, 370, 394, 467, 468, 518, 520, 595, 613,
 636, 645, 696, 781, 806–808, 847, 853
- goal of motion verb 23, 25, 130, 241, 280,
 467, 468, 595, 613
- H**
- habitual 40, 41, 171, 218, 221, 223, 224, 239,
 252, 278, 292, 333, 340, 365, 396, 457, 515,
 604, 608, 633, 636, 729, 732, 888, 912
- head 17, 36–38, 44, 118–120, 172–174, 179,
 180, 212–214, 226, 227, 239, 253, 254, 258,
 275, 284, 326, 327, 342, 343, 368, 371, 372,
 374, 426, 428, 432–434, 465, 466, 503, 504,
 522–525, 545, 546, 555, 611, 612, 617, 691,
 697–702, 708, 796, 802, 828, 834, 852, 854,
 875, 882, 900, 904, 922–924
- head-marking 691, 697
- I**
- imperative 39, 125, 127, 155, 169, 176, 183, 186,
 215, 216, 218, 223, 224, 239, 240, 251, 255,
 277, 300, 302, 328–330, 337, 340, 366, 369,
 376, 382, 425, 435–439, 443, 455–459, 462,
 463, 477, 505, 506, 513, 514, 516, 521, 526,
 547, 549–551, 565, 604–606, 621, 633–635,

- 656, 710, 715, 720–722, 818, 840, 845, 859, 871, 874, 875, 888, 906, 917, 921, 931
- imperfective 92, 126, 144, 148, 155, 186, 241, 279, 382, 457, 459, 477, 513–516, 518, 520, 521, 526, 549, 851, 871, 912, 914, 915, 931
- impossibility (Turkic verbal suffix) 366, 376, 425
- inalienable possession 212, 214, 326, 631, 635
- inchoative 767, 913, 916
- incorporation 7, 125, 144, 145, 156
- indefinite 21, 68, 73, 115, 116, 119, 139, 141, 142, 152, 155, 174, 179, 182, 186, 253, 254, 274, 275, 302, 382, 390, 418, 419, 452, 477, 502, 504, 519, 522, 525, 545, 565, 588, 598, 599, 612, 613, 621, 631, 642, 703, 704, 708, 709, 723, 770, 776, 777, 818, 862, 868, 888, 890, 903, 905
- indefinite pronoun 418, 419, 862
- indefinite marker 119, 174, 182, 502, 709
- indefinite suffix 152, 275, 504, 519
- indefiniteness 114–116, 137, 152, 273–275, 389, 390, 452, 478, 501, 502, 543–546, 584, 585, 588, 903
- index 18, 45, 163, 187, 241, 276, 286, 302, 369
- indexing 18, 125, 283–286, 552, 794, 917
- indirect object 23, 221, 333, 340, 370, 392, 464, 467, 547, 553, 591, 592, 596, 597, 613, 693, 696, 793, 853, 854, 902, 921
- infinitive 40, 50, 51, 74, 111, 112, 124, 144, 215, 216, 218, 220, 273, 276, 284, 294, 328–330, 336, 337, 382, 506, 512, 548, 565, 606, 610, 621, 688, 714, 718, 725, 733, 734, 737, 752, 754, 756, 805, 818, 828, 873, 888, 911, 912, 917, 926–928, 930–933, 950
- inflection 7, 15, 42, 44, 166–168, 210, 212, 215, 216, 240, 249, 250, 275, 324, 328, 335, 390, 417, 497, 500–502, 504, 505, 509, 514, 515, 526, 633–635, 642, 648, 869, 870, 897, 900–902, 904, 917
- instrumental 68, 122, 364, 372, 382, 396, 418, 419, 443, 594, 596, 636, 835, 865, 919
- interrogative 116, 301, 462, 517, 630, 631, 689, 694, 838, 839, 909–911, 921
- intonation 222, 230, 338, 346, 367, 426, 433, 450, 451, 492, 516, 521, 527, 647, 687, 796
- intransitivity 125, 131, 132, 144, 216, 217, 220, 282, 286, 293, 301, 329, 330, 331, 332, 335, 336, 339, 394, 395, 505, 506, 507, 508, 510, 511, 526, 549, 550, 552, 632, 687, 689, 690, 711, 727, 728, 729, 739, 740, 741, 742, 743, 765, 766, 768, 769, 773, 778, 781, 818, 841, 847, 849, 851, 852, 853, 872, 874
- Iranicization (of vowel inventory) 363, 399, 411, 428, 433
- irrealis verb form 40, 127, 218, 223, 224, 226, 227, 276, 291, 293, 332, 333, 341, 343, 344, 390, 391, 439, 509, 511–513, 518, 521, 523, 525, 526, 604, 608, 611, 621, 633–635, 648–650, 656, 729–731, 743–745, 749, 751, 818, 930, see also *realis*
- isophone (Intra-Turkic, areal) 359, 360, 416
- J**
- Jews 162, 183, 194, 195, 244, 247, 252, 253, 264, 305, 306, 309–311, 327, 351, 352, 482, 484, 485, 494, 572
- K**
- ki* (complementizer, particle, of Iranian origin) 148–150, 226, 368, 374, 375, 418, 433–435, 617, 724, 740, 744, 747, 795–805, 809–811, 924, 925
- kinship (terms) 212, 214, 325, 326, 418, 689
- L**
- labialization 288, 362, 411, 412, 493
- language academy (National Iranian) 577, 579
- language change 4, 480, 933
- language choice 303, 621, 623
- language contact 3–8, 9, 11–13, 14, 16, 18, 20–22, 25, 26, 31, 32, 42, 108, 121, 150, 181–184, 196–198, 214, 226, 228, 237, 240, 259, 260, 298, 300, 312, 344, 345, 358, 361–363, 366, 368, 370, 377, 385, 386, 388, 389, 395, 407, 410–413, 423, 426, 427, 436, 442, 485, 492, 509, 524, 526, 534, 639, 663, 667, 679, 806, 825–827, 884–886, 895–897, 903, 920, 921, 923
- language documentation 295, 860
- language endangerment → endangered language
- language loss 860, 862
- language shift 447, 624
- laryngeal 9, 33, 198, 200, 201, 238, 312, 314, 319, 387, 488, 489, 676
- laryngeal settings 200
- lenition 12, 138, 238, 271, 333, 386, 540, 541, 630
- lexicon 10, 15, 27, 48, 100, 150, 151, 181, 184, 196, 199, 202, 228, 259, 265, 271, 298, 302, 304, 305, 311, 345, 351, 358, 359, 376, 377, 396, 399, 407, 409, 414, 439, 446, 457, 478, 479, 485, 526, 578, 588, 627, 628, 657, 667, 672, 676, 820, 847, 872, 894, 951

light verb 22, 125, 130, 145, 148, 229, 273, 293, 345, 377, 464, 465, 509, 514, 520, 548, 602, 632, 635, 737, 755, 756, 765, 768, 769, 818, 872, see also complex predicate
light verb construction 125, 130, 145, 464, 465, 514, 548, 602, 765, 768, 769, 818

lingua franca 573, 575, 623, 625

linguistic area 4, 27, 28, 822, 840, see also Sprachbund, convergence zone

loanwords 10, 92, 100, 114, 138, 151, 181, 182, 196, 228, 238, 259, 260, 290, 314, 325, 345, 361, 362, 386, 387, 389, 396, 409, 490, 492, 495, 500, 502, 526, 541, 581, 583–585, 587, 588, 668, 686

locative 40, 43, 50, 79, 336, 364, 372, 377, 382, 390, 391, 415, 418, 429, 430, 443, 517, 636, 696, 705–707, 716, 718, 733–735, 737, 760, 805–808, 818, 840, 843, 844, 846, 849, 919, 929

M

macro-areas 4, 5, 28

majhul (long mid) vowels 447

masculine (gender) 16, 92, 114–116, 119, 137, 141, 155, 167, 180, 186, 249, 263, 291, 302, 321, 324, 325, 501, 544, 545, 548, 550, 552, 555, 557, 565, 587, 689, 701, 818, 819, 826, 900–904, 918, 931, 947

metathesis 143, 364, 412, 434, 486

middle voice 847, 859, 860

migration 65, 150, 160, 194, 195, 354, 398, 482, 863, 893, 895, 933

minority language 3, 163, 303, 357, 401, 406 578, 579, 825

mirative 457, 458, 633–635, 653, 656, 846

mixed language 625

modality 149, 155, 276, 375, 423, 435, 436, 567, 604, 633, 636, 861, 890, 897, 917, 919, 930, 933

mood 127, 138, 218, 420, 425, 436, 438, 439, 456, 514, 549, 632, 637, 642, 648, 840, 871, 872, 917, 920

morphology 7, 15–17, 29, 35, 51, 68, 69, 72, 74, 77–80, 83–86, 88, 90, 91, 99, 108, 114, 124, 143, 145, 150, 166, 168, 187, 196, 207, 215, 228, 235, 249, 250, 264, 267, 273, 276, 291, 300, 303, 321, 328, 345, 358, 359, 364, 365, 368, 377, 388, 396, 407, 417, 418, 420, 447, 451, 452, 455, 465, 496, 504, 505, 526, 533, 543, 585, 588, 602, 632, 634, 656, 657, 684–687, 690–692, 694, 820, 840, 846, 860, 861, 867, 871, 873, 876, 885, 897, 911, 916, 932

Mughal (Empire) 574–576, 621

multilingualism 3, 21, 196, 237, 354, 401, 405,

406, 409, 444, 538, 933, see also bilingualism

multiple relativization 882, 883

multi-purpose subordinator 857, 859, 860

N

nasal 66, 72, 164, 247, 271, 359, 364, 409, 415, 418, 448, 540–542, 581, 583, 590, 629, 719, 833, 865, 898, 899

nasalization 449

national language 570, 574, 577, 579

negation 7, 92, 125–127, 129, 138, 142, 144, 145, 148, 155, 175, 186, 224, 255, 263, 276–278, 281, 302, 366, 382, 425, 443, 462, 463, 526, 542, 549, 551, 565, 584, 585, 604, 621, 627, 630, 636, 638, 650, 710, 715, 834, 837, 843, 850, 859, 878, 885, 889, 897, 917, 919, 931, 951

neuter (gender) 900–904, 906–908, 918, 931

nominalization 371, 432, 922, 923, 927, 929, 931

nominative (case) 15, 68, 73, 92, 217, 218, 233, 240, 323, 330, 349, 389, 394, 419, 501, 531, 692, 826, 868, 870, 872, 876, 888, 902, 903, 919, 926, 931, 948

noun class 502, 589, 903

noun phrase (NP) 17, 22, 36, 37, 116–118, 120,

121, 139, 147, 239, 273, 275, 280, 281, 286, 292, 394, 428, 452, 453, 465, 467, 468, 507, 518, 524, 545–547, 551, 553, 566, 586, 601, 611, 612, 614, 631, 639, 642, 648, 690, 691, 763, 773–777, 805, 806, 818, 832, 834, 849, 854, 856, 859, 867, 868, 872, 875–880, 882, 883, 897, 900, 902–904, 907, 908, 918, 929

number 15, 116, 118, 120, 125, 139, 141, 207, 239, 249, 254, 273, 274, 276, 291, 293, 294, 321, 330, 335, 337, 389, 451, 505, 510, 543, 544, 546, 547, 552, 557, 584, 585, 587, 594, 603, 623, 631, 632, 634, 636, 637, 642, 694, 743, 777, 859, 867, 871, 872, 900, 901, 904, 906, 909, 911, 924

number agreement 116, 291, 294, 584, 603

O

object 5, 18, 21–24, 26, 43, 44, 115, 125, 131, 132, 168, 169, 209, 221, 223, 240, 241, 250, 254, 264, 280, 281, 283, 286, 294, 301, 331, 333–335, 340, 345, 368, 370, 393–395, 419, 426, 435, 451, 452, 455, 462, 464, 465, 467, 477, 500, 507–511, 514, 515, 518–520, 524, 544, 547, 548, 550, 553, 566, 590–593, 595, 597–599, 602, 613, 614, 621, 632, 635, 636

- 638, 650, 689, 691–693, 696, 733, 770, 786–789, 791, 793, 805, 826, 827, 841, 844, 847, 850, 851, 867, 876, 877, 902, 903, 905, 906, 921, see also direct object, indirect object
object agreement 520, 727
- oblique 15, 17, 37, 44, 114–117, 119, 120, 131, 132, 135, 137, 147, 148, 155, 233, 239, 291, 293, 301, 302, 323, 349, 370, 382, 389, 393, 394, 426, 427, 501, 503, 505, 507, 515, 531, 533, 544, 547, 553, 554, 565, 586, 650, 689–691, 694–702, 727, 755, 756, 763, 764, 770, 772–783, 789, 791, 793, 794, 806, 818, 819, 866, 868, 870, 875, 876, 888, 948
- oblique case 17, 37, 44, 114–116, 119, 120, 131, 132, 135, 137, 147, 239, 291, 293, 301, 370, 394, 426, 533, 544, 547, 553, 689–691, 696–700, 755, 756, 763, 764, 770, 772, 774, 778, 780–782, 791, 793, 794, 795, 866, 868
- official language 3, 162, 196, 237, 244, 357, 398, 405, 485, 538, 570, 572, 574, 575, 579
- onset, syllable 112, 272, 318, 319, 341, 361, 448, 449, 833, 899
- optative 176, 255, 376, 382, 383, 436–439, 443, 608, 689, 721, 723–725, 730, 731, 753, 819, 845, 859, 917, 920, 931
- orthography 109–111, 121, 152, 288, 289, 311, 365, 447, 478, 539, 542, 627, 861, 863, 894, 897
- Ottoman (Empire) 1, 383, 574
- Ottoman culture 259
- Ottoman government 309, 355
- OV (word order, including SOV) 5, 21–24, 130, 241, 273, 281, 339, 428, 467, 614, 627, 776–779, 807, 827, 852, 876, 921, 933
- P**
- palatalization 12, 33, 200, 238, 271, 314, 388, 396, 413, 417, 676, 717, 866
- participle 40–42, 50, 69, 78, 86, 92, 137, 155, 168–171, 183, 186, 215–217, 220, 240, 251, 252, 263, 278, 279, 328–330, 332, 334, 335, 341, 366, 367, 372–374, 382, 392, 429, 432, 435, 458, 460, 477, 506, 510, 511, 513, 517, 519, 548, 552, 557, 565, 602, 606–610, 621, 648, 650, 685, 718, 738, 741, 743, 760, 761, 767, 806, 819, 850, 851, 859, 866, 873, 888
- passive 169, 219, 220, 251, 255, 263, 273, 328, 332, 366, 405, 406, 513, 548, 603, 608–610, 632, 689, 767–769, 819, 832, 873, 916, 926, 931
- past (tense) 7, 115, 125–127, 131, 135, 146–148, 171, 241, 252, 276, 281, 283, 284, 286, 290, 293, 547, 552, 553, 603, 606, 608, 621, 727, 773, 827, 881, 912, 914, 926, 927
- patient 258, 688, 689, 696, 721, 727, 773–782, 784, 785, 791
- perfect 41, 42, 44, 92, 140, 145, 148, 149, 155, 167–169, 171, 175, 183, 216, 217, 220, 221, 240, 250–252, 255, 279, 302, 334–336, 341, 351, 366, 367, 369, 392, 394, 395, 424, 456, 458–460, 477, 511, 515, 519, 526, 532, 552, 557, 560, 565, 585, 606–609, 611, 633–635, 645, 648, 650, 656, 663, 685, 710, 738–743, 747, 749–751, 763, 790, 791, 819, 826, 843, 845, 912, 931
- perfective 6, 41, 42, 44, 186, 216, 217, 220, 240, 241, 263, 331, 333, 335, 336, 382, 394, 395, 443, 455, 513–515, 520, 526, 552, 851, 859, 871, 912–914, 917
- personal pronoun → pronoun, personal
- person marking 42, 125, 126, 129, 145–147, 241, 277–279, 299, 304, 330, 335, 455, 457–461, 466, 840, 841, 843
- Persophonie 622
- pharyngeal 10, 11, 32, 33, 113, 114, 139, 156, 198–201, 271, 288, 312–316, 345, 361, 363, 386, 387, 410, 485, 486, 488–490, 540, 541, 583, 628, 629
- pharyngealization 9, 11, 33, 113, 114, 150, 201, 202, 315, 316, 386, 387, 486, 490, 826
- phoneme 9, 11, 32, 33, 109, 110, 113, 139, 164, 165, 182, 197, 198, 235, 247, 248, 271, 272, 287, 312, 316, 316–318, 360–362, 387, 408–411, 539, 540–542, 581–583, 627–630, 676–679, 809, 833, 898, 899
- phonemicized allophones 197, 312, 539, see also allophone
- phrasal verbs 182, 229, 345, 509, 602
- pluperfect 155, 279, 460, 552, 607, 618, 843, 845
- plural 11, 15, 16, 35, 73, 74, 83, 85, 92, 114, 116, 117, 119, 131, 139, 141, 142, 147, 155, 166–169, 186, 213, 214, 239, 249, 263, 291, 299, 302, 321, 323–326, 359, 361, 364–366, 369, 382, 414, 416, 417, 420–423, 429, 443, 451–453, 455, 477, 486, 497, 498, 501, 502, 515, 542, 544–550, 552, 557, 565, 584–589, 603, 605, 621, 631, 633, 637, 642, 645, 656, 688–692, 700, 703, 721, 774, 781, 818, 819, 834, 842, 843, 846–848, 854, 859, 867, 868, 872, 875, 886–888, 900–902, 904, 907, 911, 931
- politeness, system of 578, 590, 603
- possessive 11, 14, 18, 40, 85, 87, 151, 211, 212, 239, 273, 287, 302, 323, 324, 359, 364, 367,

- 369, 372, 373, 377, 382, 393, 416–421, 426–428, 431, 432, 443, 512, 548, 551, 584, 591, 631, 637, 641–643, 645, 679, 688, 689, 691, 692, 699, 702, 709, 710, 763, 764, 770, 778, 783–786, 789, 806, 819, 827, 834, 835, 849, 859, 875, 878, 885, 888, 907, 908, 927, 931
- possessor 85, 115, 117–119, 132, 133, 148, 281, 282, 284, 286, 371, 392, 393, 418, 427, 547, 548, 593, 642, 691, 696–700, 763, 764, 784, 843, 847–849, 867, 907, 908
- post-constituent negation 627, 638
- potential 278, 365, 366, 370, 373, 382, 425, 432, 439, 608, 745, 826, 842, 843, 848, 849, 859, 885 see also impossibility (Turkic verbal suffix)
- pragmatics 478, 479, 622
- predicative 18, 44, 137, 139, 393, 759, 763, 764, 821, 851
- preposed relative clauses 806
- preposition 39, 68, 112, 122, 123, 129, 131, 134, 135, 139, 142, 143, 253, 254, 257, 280, 283, 284, 330, 336, 340, 370, 393, 428, 454, 465, 466, 468, 505, 515, 520, 525, 543, 545, 553, 589, 595, 596, 601, 630, 636, 641, 642, 705, 706, 718, 794, 805, 875, 919, 923, 927, 930
- present (tense) 51, 96, 115, 125, 140, 143, 146, 147, 170, 241, 252, 277, 278, 281, 293, 298, 301, 394, 422, 544, 547, 549, 552, 553, 604–606, 621, 687, 688, 716–720, 729, 734, 737, 826, 827, 840, 912, 920, 925
- preverb 144, 464, 548, 591, 604, 688, 712–716, 728–731, 765, 819, 840, 844–846, 850, 859, 871, 872, 888
- pro-drop 590, 613, 876
- progressive 40, 41, 51, 74, 126, 140, 145, 170, 218, 220, 221, 239, 252, 278, 336, 337, 341, 364, 365, 391, 396, 414, 415, 421, 457, 460, 461, 511, 512, 515, 518, 603, 609, 633, 688, 718, 729, 732–738, 790, 819, 915, 916
- progressive present 170, 252, 396, 421, 511
- proleptic 880, 881
- pronoun, personal 117, 119, 167, 172, 175, 209, 250, 273, 274, 280, 291, 367, 368, 408, 419, 420, 426, 427, 498, 515, 526, 546, 589, 631, 869, 905, 935, 948, see also demonstrative, indefinite, interrogative, relative, resumptive pronoun, clitic 18, 22, 240, 273, 274, 281–283, 286, 330, 368, 369, 464, 526, 541, 547, 552, 553, 556, 584, 591, 593, 907, 911, see also copied Iranian enclitic pronouns
- Q**
- quantifier 630, 708
- question 50, 128, 163, 166, 186, 187, 259, 299, 367, 425, 426, 588, 599, 614, 631, 650, 656, 759, 856, 884, 885, 906, 910, 911, 915, 924, 931
- R**
- realis verb form 218, 227, 344, 391, 512, 513, 515, 516, 525, 630, 633–635, 652, 653, 656, see also irrealis
- reciprocal 117, 377, 382, 859, 873
- reconstruction (historical) 21, 48, 207, 209, 486, 933
- reflexive 117, 119, 131, 148, 151, 155, 263, 300, 302, 367, 419, 420, 455, 546, 547, 565, 593, 594, 621, 631, 656, 709, 710, 873
- relative pronoun 173, 179, 181, 631, 648, 882, 883, 908, 922, 923
- relativizer 118, 156, 186, 303, 909, 922–924, 931
- resumptive 524, 553, 924
- resumptive pronoun 524, 924
- re-syllabification 166
- rhotacism 364, 412
- root 47, 93, 113, 287, 328, 361, 377, 412, 414, 455, 482, 634, 643, 688, 711–714, 718, 720, 722, 733, 742, 743, 754, 755, 758, 765, 789, 840, 844, 846, 849, 850, 871, 872, 885, 917, 923
- rounding (vowels) 13, 14, 15, 111, 138, 204, 238, 272, 289, 300, 362, 363, 396, 410, 411, 417, 420, 487, 582, 678, 947, 949
- S**
- Sassanian (Empire) 243, 625, 666
- semantics 7, 286, 293, 304, 457, 566, 631, 632, 639, 657, 796, 826, 829, 871
- sociolinguistics 158, 162, 163, 188, 196, 246, 311, 357, 398, 446, 478, 485, 537, 569, 577, 624, 821, 825, 892, 893, 933
- sound shift 50, 61, 248
sound shift r > ġ 248
- SOV (word order) → OV
- specifier 368, 477
- Sprachbund 27, 891, see also linguistic area
- stranding 123
- stress 15, 118, 206, 207, 222, 230, 249, 300, 320, 321, 338, 346, 450–452, 477, 492–495, 498, 516, 527, 539–544, 549, 584, 585, 605, 607, 622, 630, 642, 684–686, 688, 715, 720, 738–740, 742, 743, 833, 835, 866, 867, 889, 911

- subject 51, 74, 111, 113, 115, 117, 119, 120, 125, 128, 131, 132, 137, 141, 146–148, 151, 166, 172, 179, 182, 217, 218, 241, 250, 281, 286, 292, 293, 330–332, 334, 335, 338, 342, 368, 374, 393–395, 414, 417, 464, 466, 477, 481, 498, 505, 507, 510, 511, 515, 520, 521, 547, 549, 550, 552, 587, 590–592, 601, 613, 614, 631, 632, 636, 645, 649, 650, 678, 695, 696, 752, 755, 757, 779, 781, 783, 790, 826, 827, 841, 843, 847–849, 851–853, 863, 876–878, 902, 903, 905, 911, 916, 922–924, 926
 subject marking 149, 464, 511
 subject, non-canonical 132, 150, 286, 287, 293, 843, 848, 853
 subject relativization 286, 828
 subjunctive 40, 41, 50, 51, 69, 70, 74, 112, 125–129, 155, 239, 277–279, 284, 291, 292, 302, 376, 390, 436–439, 456–458, 462, 463, 477, 542, 549, 551, 552, 565, 604, 606, 610, 611, 615, 617, 619–621, 678, 685, 688, 710, 718, 720–725, 727, 730–732, 741–753, 757, 758, 762, 803, 819, 845, 859, 871, 872, 874, 875, 888, 917, 920
 subordination 25, 26, 241, 429, 465, 468, 607, 620, 630, 636, 647, 648, 795, 796, 805, 807, 880, 922, 923, 926
 suffix 7, 8, 11, 14–17, 35, 44, 87, 114, 116, 118, 124, 137, 139, 147, 148, 152, 166–168, 174, 179–181, 183, 210, 212–214, 226, 227, 233, 239, 240, 248, 249, 251, 252, 274, 275, 279, 283, 284, 286, 291, 292, 300, 322, 326, 327, 330, 333, 336, 349, 364, 366, 369, 372, 373, 376, 377, 389, 390, 393, 414, 416, 418–420, 422, 423, 425, 427, 428, 436, 451–453, 456, 458–460, 495, 499, 501, 502, 504, 513, 515, 519, 520, 524, 526, 531, 543, 550–554, 584–586, 588, 589, 608, 615, 624, 630–633, 643, 644, 656, 687, 690–692, 699, 716, 719, 729–731, 733, 765, 826, 827, 834, 836, 837, 840–842, 844, 845, 847, 849–851, 859, 868, 872, 874, 886, 887, 900–902, 906, 907, 911, 917, 918, 929
 superlative (adjective) 370, 419, 426, 584, 930
 SVO (word order) → VO
 syllable 48, 90, 108, 111–113, 143, 165, 202, 203, 205, 249, 272, 317–321, 361, 363, 365, 367, 387, 397, 410, 412, 426, 433, 449, 450, 454, 477, 488, 489, 492–495, 539, 542, 543, 549, 573, 583, 584, 629, 630, 686, 720, 733, 796, 833, 866, 867, 899–901
 syntax 7, 16, 22, 24, 26, 27, 38, 44, 71, 76, 78, 80, 84, 87, 89, 91, 120, 130, 131, 148, 150–152, 172, 176, 187, 222, 225, 236, 241, 252, 256, 260, 282, 284, 298, 301, 326, 338, 341, 345, 358, 367, 371, 391, 392, 394, 396, 399, 429, 430, 433, 444, 447, 464, 465, 516, 521, 526, 580, 602, 611, 615, 621, 622, 627, 636, 638, 642, 657, 827, 860, 890, 895, 897, 916–918, 931
- T**
- topic 275, 282, 452, 466, 543, 659, 665, 744, 836, 837, 888, 900
 topicalization 599
 transitional zone 5, 355, 400, 807
 transitive 44, 115, 125, 131, 132, 135, 216–218, 220, 221, 241, 276, 281–283, 285, 286, 293, 302, 329–332, 334, 336, 339, 394, 395, 477, 505–508, 510, 511, 544, 548–550, 552, 553, 595, 599, 632, 687, 689, 690, 711, 721, 727–731, 739–743, 763, 765, 768–770, 778, 791, 826, 841, 842, 847–851, 853, 859, 872, 874, 906
 transitivity 126, 220, 329, 331, 552, 826, 917
 Turkoman (of Iraq) 244, 246, 259, 311, 345
- U**
- unaccusative 144, 220, 508, 510, 511, 526, 767, 769, 770
 unergative 507, 508
- V**
- valency 765, 840, 843, 846, 847, 849, 860
 valency-changing 846
 verb 6, 7, 13, 18–26, 40, 41, 50, 51, 70, 74–76, 80, 84, 86, 91, 112, 115, 116, 122, 125–132, 134–136, 139, 144, 145, 148, 149, 159, 166, 170–172, 175, 176, 181, 182, 223, 226, 227, 229, 234, 240, 241, 244, 249, 252, 264, 271, 273, 276–281, 283, 286, 292–294, 297–299, 328, 331–335, 337, 340, 341, 345, 363, 365, 366, 368–371, 373–377, 390–392, 394, 395, 414, 420, 422, 425–433, 435–439, 446, 450, 455–470, 477, 498, 505–512, 514–516, 518–520, 525, 526, 548–553, 560, 563, 587, 590, 592, 595, 599, 601–603, 605–607, 609, 610, 613, 615, 617, 618, 627, 630–638, 642, 644, 645, 647–653, 656, 684, 687–690, 710–714, 716, 718, 720, 723, 725, 727, 729–731, 735–738, 740, 742–748, 750–752, 754–757, 759, 762, 763, 765–769, 773–778, 786, 787, 789, 794, 803, 805, 818, 820, 821, 826–828, 832–837, 839–841, 843–852, 854–856, 861, 866,

- 871–874, 876, 878, 881, 882, 885–887, 890, 905, 911–913, 915–917, 920–924, 947, 949
- verbal modifier 170, 183, 252, 257, 260
- verb phrase (VP) 139, 281–283, 302, 468, 469, 602, 627, 635, 636, 645, 787, 876
- verb stem 7, 112, 114, 278, 279, 363, 365, 414, 422, 462, 505, 506, 548, 633, 637, 648, 690, 710, 711, 716, 740, 911
- VO (word order, including SVO) 5, 21–23, 28, 244, 339, 807, 827, 921, 933
- voice (verbal) 220, 632, 637, 767, 847, 859, 860, 899, 911, 916, see also middle voice
- voice, voiced (phonation) 10–12, 50, 61, 66, 77, 79, 81, 88, 90, 96, 98, 103, 113, 137, 164, 165, 200, 201, 238, 247, 271, 312–315, 360, 363, 386, 387, 407–410, 412, 414, 448, 449, 488, 489, 540, 541, 543, 564, 581, 582, 627, 628, 676, 677, 833, 866, 898, 899, 947, 949
- voiceless (phonation) 10, 50, 79, 108, 113, 150, 164, 238, 247, 271, 288, 407, 448, 449, 488, 540, 541, 581, 583, 627, 676, 677, 825, 833, 898–900
- volitive/volitional 41, 218, 926
- vowel 12–15, 33, 34, 39, 41, 61, 74, 81, 83, 85, 90, 109–112, 115, 122, 126, 129, 137–139, 143, 150, 165–168, 182, 197, 199, 200, 202–206, 218, 219, 238, 248–250, 1, 272, 277, 279, 288–290, 300, 312, 314, 316–320, 326, 328–330, 360–363, 365, 386–388, 396, 408–414, 417, 420–422, 433, 447–452, 454, 455, 459, 465, 468, 477, 485, 487, 488, 491–495, 498, 500, 518, 539–545, 547, 549, 550, 557, 580–584, 591, 605, 628–630, 633, 637, 643, 656, 676–683, 687, 690, 692, 708, 715, 717–721, 733, 737, 739, 757, 832, 833, 839, 865–867, 869, 898–900, 902, 904, 907, 910, 911, 947, 949
- vowel harmony 8, 363, 396, 715, 910, 911
- W**
- Wackernagel rule 787
- wh-word 910, 911
- word order 5, 21–24, 27–29, 130, 134, 156, 176, 187, 223, 241, 264, 273, 280, 370, 371, 428, 518, 612–614, 622, 688, 777, 792, 793, 795, 796, 802, 806–808, 827, 852, 861, 876, 878, 890, 897, 921
- Y**
- Yezidi 140, 237, 246
- Z**
- Zagros-d 413
- ?**
- ?imāla 165, 248