

### 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.<sup>1</sup> 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 Bājaḷānī (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.

---

<sup>1</sup> 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.

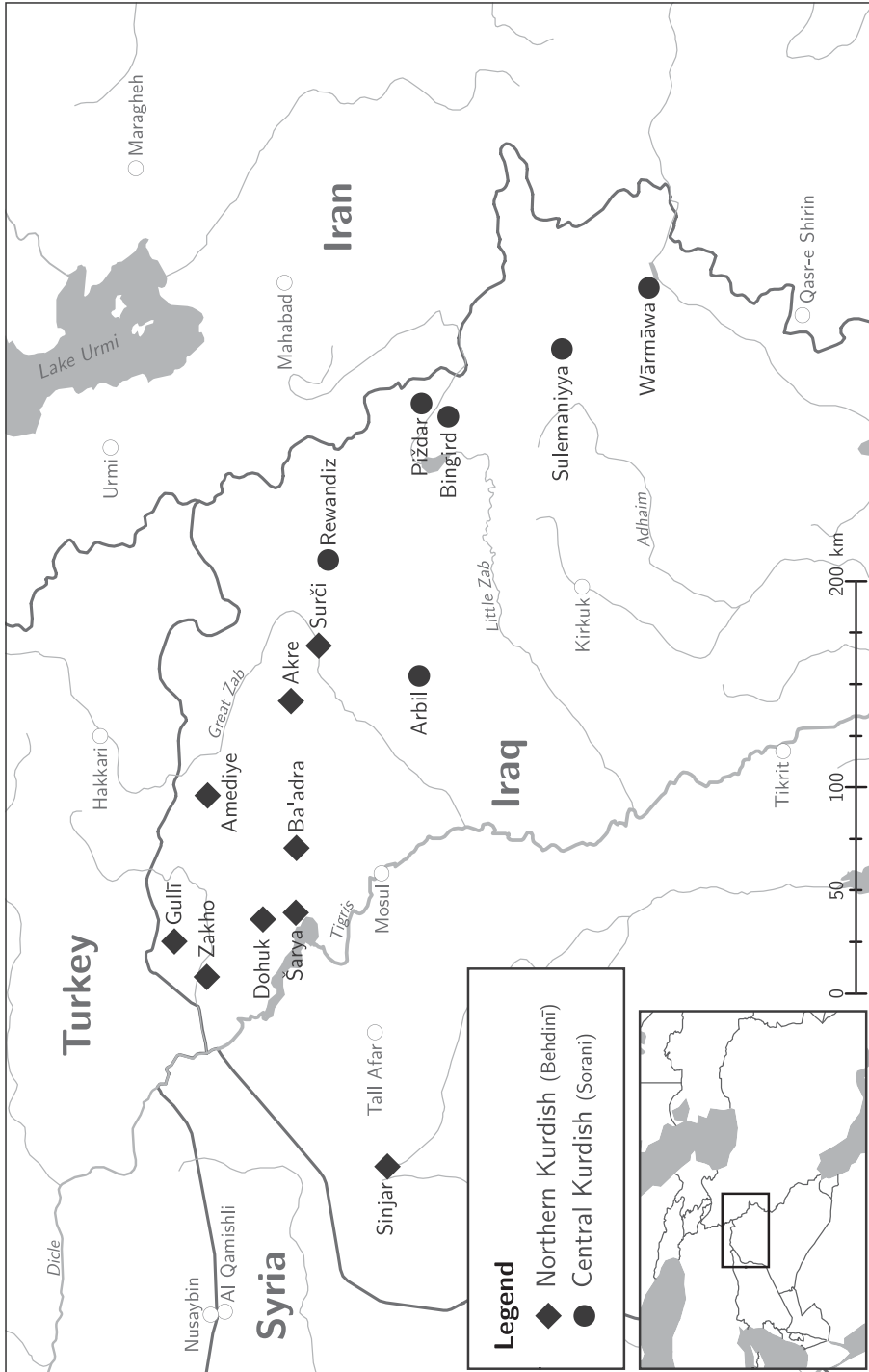


Figure 1: Approximate locations of Northern and Central Kurdish varieties mentioned in this chapter

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 Ardalans 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 Sulemaniyya, 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 Ezidî 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 Öpengin (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 Öpengin (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			ʧ ʤ					
Nasal	m	n			ŋ			
Fricative	f v	s z	ʃ ʒ		x ɣ		ħ h	h
Tap/Trill		r / r						
Approximant								
Lateral		l	ɭ					
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: *birīndar* [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.jəm] ‘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 [ʧ, ʤ] respectively, and the fronting of the original affricates to [tɕ, dʒ] (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.

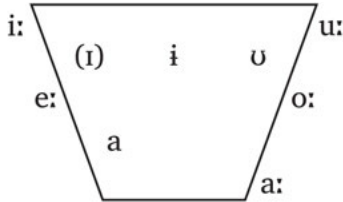


Figure 2: Vowels of Sorani

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 [ɪmtɪhān] ‘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 [hɛ.jə] ‘there is’, and low central [a] adjacent to pharyngealized [ʂ], as in [ʂəst] ‘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’, *tfeŋ* ‘rifle’, *xrāp* ‘bad’. However, these complex onsets can generally be analysed as resulting from elision of underlying [j], 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 (Öpengin and Haig 2014; Haig and Öpengin 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: *kuştin* ‘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-poselector 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 *hāt wāsīn* ‘hang up’, with the particle *hāt* 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-suffixed 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 = $\bar{i}$  (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 *- $\bar{e}k$*  (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 xoš $\bar{i}$ =y ew-e*  
 from pleasure= $\bar{E}Z$  that= $\bar{D}EM$   
 ‘from the pleasure of that’ (MacKenzie 1961: 62)
- (2) *tūtik- $\bar{e}k$ = $\bar{i}$  pičkole*  
 dog= $\bar{I}NDF$ = $\bar{E}Z$  small  
 ‘a little dog’ (MacKenzie 1961: 63)
- (3) *xānū=y ēme*  
 house= $\bar{E}Z$  1 $\bar{P}L$   
 ‘our house’ (MacKenzie 1961: 63)
- (4) *kič- $\bar{e}k$ = $\bar{i}$  đwān= $\bar{i}$  čwārde sāl*  
 girl= $\bar{I}NDF$ = $\bar{E}Z$  beautiful= $\bar{E}Z$  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 *- $\bar{e}k$* , 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= $\bar{C}EZ$  pretty= $\bar{D}EF$   
 ‘the pretty girl’ (McCarus 2009: 613)

- (6) *kilk=e ker-ĕk*  
 tail=CEZ donkey-INDF  
 ‘a donkey’s tail’ (MacKenzie 1961: 65)
- (7) *ew xĕwet=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>ĉūn / řōīštin</i>	<i>ĉū- / řō-</i>	<i>ĉ-</i>
come	<i>hātin</i>	<i>hāt-</i>	<i>(h)ĕ-</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>tĕ-</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 irrealis 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 Öpengin (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	<i>-im</i>	<i>-im</i>	<i>=im</i>
	2	<i>-ī(t), -y(t)</i>	<i>-ī(t)</i>	<i>=ī</i>
	3	<i>-ē(t)<sup>2</sup></i>	<i>-∅</i>	<i>=e</i>
PL	1	<i>-īn</i>	<i>-īn</i>	<i>=īn</i>
	2	<i>-in</i>	<i>-in</i>	<i>=in</i>
	3	<i>-in</i>	<i>-in</i>	<i>=in</i>

### 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)	<i>e-</i>	<i>e-tē-m</i> ‘I say, am saying’
negated (NEG)	<i>nā-</i>	<i>nā-tē-m</i> ‘I don’t say’
subjunctive (SUBJ)	<i>bi-</i>	<i>bi-tē-m</i> ‘that I say’
subjunctive negated (SUBJ.NEG)	<i>ne-</i>	<i>ne-tē-m</i> ‘that I not say’
imperative (IMP)	<i>bi-</i>	<i>bi-tē</i> (2SG), <i>bi-tē-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)	<i>me-</i>	<i>me-tē, me-tē-n</i>

<sup>2</sup> 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:<sup>3</sup>

- (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=kuž-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 unpre-fixed, 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).

<sup>3</sup> 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 *hātin* ‘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  $=\text{(y)}e$ , 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) *hat=ī*      *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 Sangesari (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	$=\text{(i)}m$	$=\text{(i)}t$	$=\bar{i} / =y$
PL	$=mān$	$=tān$	$=yān$

Within the languages conventionally assigned to ‘Kurdish’ (Haig and Öpengin 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 Surçi 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-kuřt*  
 1SG dog-DEF=1SG NEG-kill.PST(3SG)  
 ‘I didn’t kill the dog’  
 b. *min ne=m=kuřt*  
 1SG NEG=1SG=kill.PST(3SG)  
 ‘I didn’t kill (it)’  
 c. *min kuřt=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=1 PL.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=1 SG                                    SUBJ-give.PRS.2SG=ASP  
 ‘Give me back my quarry’ (MacKenzie 1962: 8)
- b. *mīn žin=im*                    *nā-wē*  
 1 SG woman=1 SG NEG-be.desirable.PRS.3SG  
 ‘I don’t want a wife’ (MacKenzie 1962: 6)
- c. *heyās ū ḥesen memend=im*            *be hīč*            *le des*  
 Heyas and Hasan Mamand=1 SG    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:<sup>4</sup>

- (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-lē* ‘say to me!’ (with an independent pronoun as complement of the preposition *be*) vs. *pē=m bi-lē!* (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)

<sup>4</sup> 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*    *eħme-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 dinye=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*    *hetstā*  
 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: *selāhedīn, ke dīnye=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’ through 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 Öpengin 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 Öpengin 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 <sup>h</sup> p b			t <sup>h</sup> t ʈ d			k <sup>h</sup> k g	q		ʔ
Fric.		v f	s z	ʃ ʒ			x ɣ	ʁ	ħ ʕ <sup>5</sup>	h
Affr.					ʃ <sup>h</sup> ʒ <sup>h</sup> dʒ					
Nas.	m			n			ŋ			
Trill				r						
Flap				ɾ						
Approx.	w					j				
Lateral			l (dialectally also ɭ)							

<sup>5</sup> 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 /t̤ ʂ z/ 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 *şebr* ‘patience’, or *teyr* ‘bird’, but also occur in the native vocabulary of Behd., e. g. *şār* ‘cold’, *ṭārī* ‘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 include *tā* ‘fever’ vs. *ṭā* ‘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	(û)	ü
[u:]	ô	û	ū
[ʊ]	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.

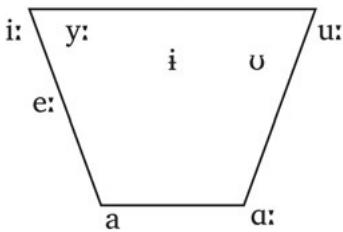


Figure 3: Vowel system of Behd., variety of Akre

Vowel length is not by itself distinctive, though [i, y, u, e, ɑ] 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 [mīro:v] ‘person’ is [mīruv], and [go:t] is [gut] ‘said’ in Zakho. This latter process is also noted for Šemzīnan 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 *suq*. 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:nifti:</i>	<i>ru:nifti:</i>	<i>ry:nifti</i>	<i>ri:nifti:</i>	<i>ri:nifti:</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: *-î* (masc. sg.) vs. *-ê* (fem. sg.), and the forms of the Ezafe: *⇒yê* (masc. sg.) vs. *⇒yâ* (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 *-î*: *min gut=e zelâm-î* 'I(OBL) said=to the man-OBL'. The plural form of the Ezafe is *-êt*, rather than the *-ê(n)* of most of NK. The paradigm of personal pronouns is identical to the rest of NK, except for the second person plural in Surçî and Akre, where the forms *engo* (Surçî, 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 (*-în* 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 Surçî, Amadiye, Akre and Dohuk the suffix is *-ît*, as in Dohuk *çê-nâ-b-ît* PRV-NEG-be.PRS-3SG 'It doesn't happen'. In Zakho and Gullî we find *-it* (*t-ç-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: *-ît* (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' *dân* '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*  
 man-INDF NEG that strength ADD FUT can.PRS-3SG  
*wan dār-ā bi dest-ē xo hīl-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.<sup>6</sup>

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) *l 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-prefixed present stem. For example, an account of earlier traditional wedding celebrations (Haig and Mustafa 2016) begins with the following:

<sup>6</sup> 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;<sup>7</sup> 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:

<sup>7</sup> 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 pičikā 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*  
 1 SG.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*  
 1 SG    in.it    NEG-COP.PRS-1 SG  
 ‘(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-1 SG    sign-INDF    such    existent=COP.3SG  
 ‘(saying) that I know there is such a sign’
- (41) *wāxtē čū hevāl ū bend-ēt*                      *xo helā-n<sup>8</sup>*  
 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?”’

<sup>8</sup> The verb is *helān* ‘leave’, which in other dialects has the infinitive *hištîn*. 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 McKenzie 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 Kermānšāh (see Fig. 4). This variety is close to 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).

**Gawṛāḡūyī**

The variety spoken only in the village of Gawṛāḡu, 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 Gawṛāḡu 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).

**Bāḡaḡānī**

This is a term loosely identified with the speech of the Bāḡaḡān 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 Bāḡaḡān 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 Bāḡaḡān are known for their religious heterodoxy.

**Šabakī**

The speakers are socially and linguistically closely aligned with the Bāḡaḡān, 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 Bāḡaḡānī (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 Tōpẓāwa near Kirkūk, Xānaqīn and Arbil (Bailey 2016: 644). The language is often referred to with the term *Mačō* (3sg present of the verb ‘say’).

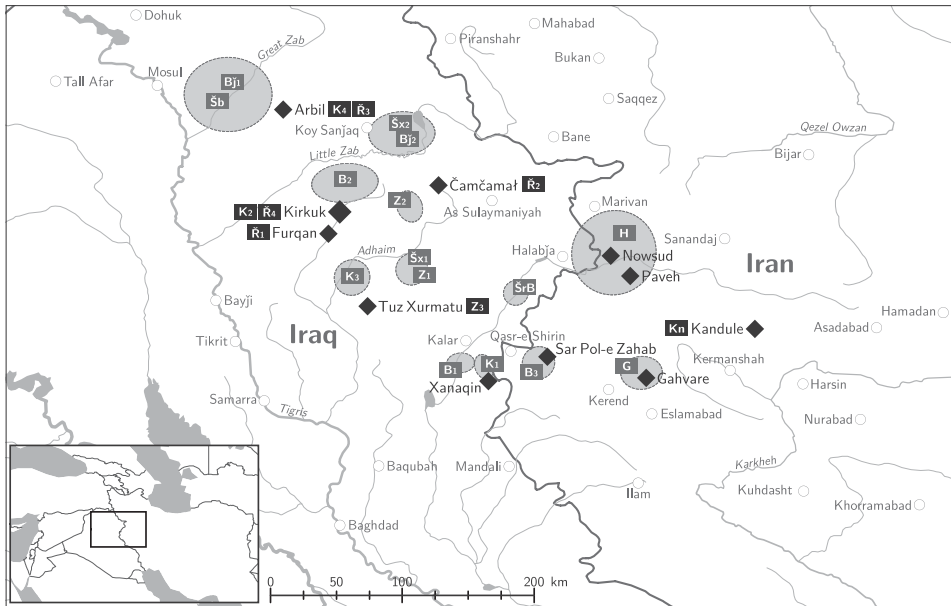


Figure 4: Names and locations of Gorani speaking localities (adapted from Bailey 2016: 9–10)

### Key to Abbreviations used in Fig. 4:

**B (Bēwyānī)** B1: Sarqizil, Bardī ‘Alī Xwārū/Žūrū, Say Mišafā, Gōrī Ginūž, Kānī Širna; B2: Bēwyānī Gawra/Bičūk and about another fifteen villages; B3: Bēwyān, Dūšamnān (their main places in this area), Sar Pol-e Zahāb, Dāraka, Qalāma, Šāy Tōtyā, Barxu Bārānī Xwārū/Žūrū, Mižūryānī Ambar/Awbar, Tangī Hamām, Sarqaḷā

**Bj (Baǰalānī)** Bj1: several villages near Al-Mawšil (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 Gawraǰū villages

**H (Hawrāmān area)** Center: Pāve and Newsūd

**K (Kākayī)** K1: Xānaqīn, Mēxās, Qarāmīn, Malā Rāhmān, Dārā, Qalāma, ‘Alī Bāpīr, Rāmazān, Tapa Čarmē, Hāǰī Mišafā; K2: Kirkūk (districts Hayy ‘Askarī, Usarā’ al-Mafqūdīn etc.); K3: around Tōpẓāwa (center of the Kākayī), ‘Alī Sarāy, Garǰa Kōyī, Ma‘fīq, Dāquq, Zaqar; K4: near Arbil: Sufāya, Wardak, Kabarlū, Tūlaban, Gazakān

**Kn (Kandūla)** three villages

**Ř (Řožbayānī)** Ř1: Furqān; Ř2: Čamčamā; Ř3: Arbīl; Ř4: Kirkūk and Laylān (near Kirkūk)

**Šb (Šabak)** Šabak villages near Al-Mawšil

**ŠrB (Šaraf Bayānī)** villages around Bamō

**Šx (Part of Šēxānī)** Šx1: villages near Qādir Karam: Qašqa, Wēla, Šawak, etc.; Šx2: Kānī Māz, Taqtaq, Sē Girdkān, Qāmīš, (Kōya (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-Mawšil); 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 Bājaḷānī (Šabakī, 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> [møjny]	<i>-ī/y</i>	<i>-im</i>
2SG	<i>m-us-ī</i>	<i>m-ōs-ī</i>	<i>me-wīn-ū</i>	<i>-ī</i>	<i>-īt, -yt</i>
3SG	<i>m-us-o</i>	<i>m-ōs-ō</i>	<i>me-wīn-o</i>	<i>-o</i>	<i>-ē(t)</i>
1PL	<i>m-us-me</i>	<i>m-ōs-mē</i>	<i>me-wīn-im</i>	<i>-mē</i>	<i>-īn</i>
2PL	<i>m-us-de</i>	<i>m-ōs-ē</i>	<i>me-wīn-de</i>	<i>-dē</i>	<i>-in</i>
3PL	<i>m-us-ā</i>	<i>m-ōs-ān</i>	<i>me-wīn-ā</i>	<i>-ān, -ānē</i>	<i>-in</i>

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 Bājaḫānī, 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 Gawraḡjūyī (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 Kurdicized variety of Gorani yet documented. A question mark indicates lack of relevant data; I have

collapsed Bailey's two columns "Bājaḷānī" and "Bājaḷānī/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) <sup>9</sup>	– ( <i>ištān</i> )	+	+	+	+	<i>hē</i> ( <i>&lt;*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). <sup>10</sup>	+	+	+	+	+	+

<sup>9</sup> According to Sara Belevi (p.c. 17. 09. 2016), a reflexive pronoun *wiž* (possible variant *viž*) is present in most 'Laki-Kermānshāhi' varieties (e. g. Bisotun, Chehr, Harsin, Pāyравand: see Fattah, 2000: 291 and Belevi, 2016: 65–66). Depending on how one defines "Southern Kurdish", this feature may not qualify. It is also attested in varieties of Lekī.

<sup>10</sup> One might conjecture on possible influence of the Persian *m̄-* 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 Belevi (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, Pāyравand). This morpheme is also typical of Laki dialects (see, for instance, Lazard, 1992: 218); see Fattah (2000: 371–372), Belevi (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) <sup>11</sup>	+	+	+	+	+	+
‘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>mangū</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

<sup>11</sup> 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 Gawraǰū 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 Kamandâ. 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-Hårsta & 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 Öpengin, 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 Öpengin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*. Bamberg: Bamberg University Press.
- Haig, Geoffrey & Ergin Öpengin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 3. 99–122.
- Haig, Geoffrey & Ergin Öpengin. 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. *ILLIC Technical Reports Series X*. Amsterdam: Institute for Language, Logic, Language, and Computation.
- MacKenzie, David. 1956. Bājalānī. *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 Gawraji, 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.
- Öpengin, Ergin. 2016. *The Mukri variety of Central Kurdish. Grammar and Texts*. Wiesbaden: Reichert.
- Öpengin, Ergin. In press. Accounting for the combination of clitic and affix person markers in Central Kurdish. In Songül Gündoğdu, Ergin Öpengin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*. Bamberg: Bamberg University Press.
- Öpengin, Ergin & Geoffrey Haig. 2014. Regional variation in Kurmanji: A preliminary classification of dialects. *Kurdish Studies* 3. 143–176.
- Samvelian, Pollet. 2007. A lexical account of Sorani Kurdish prepositions. In Stefan Muller (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.