

CHAPTER 1

BACKGROUND TO THE LANGUAGE, COMMUNITY, AND FIELDWORK

1.1 Introduction

The present work is a grammatical description of the Mukri variety of Central Kurdish accompanied by a collection of texts and a brief lexicon. More theoretical aspects of person marking and argument indexation patterns in Central Kurdish are presented in a companion volume currently being prepared for publication (Öpengin forthc.). Central Kurdish, more widely known as Sorani, is one of the major varieties of Kurdish, itself a branch of Western Iranian languages (cf. Korn 2003). Central Kurdish is spoken in Iraq and Iran by the majority of the Kurdish population in these two countries. McCarus (2009:587) estimates the number of its speakers as around 5 million, while in Lewis, Simons & Fennig (2013) the number is more specifically stated as 6,750,000. This background chapter presents the language and the speech community under study (§1.2); the position of the variety within Kurdish dialectology (§1.3); a synopsis of previous work (§1.4); a description of the fieldwork and data collection process (§1.5), and a presentation of the corpus of this study (§1.6).

1.2 Mukri Central Kurdish and its speech community

Mukri (more precisely *Mukri*) or *Mukriyani* is the indigenous name for the variety of Central Kurdish spoken in the northern half of the Central Kurdish speech area in Iran. Figure 1 shows the speech zone of Mukri within (Central) Kurdish and with respect to its contact languages. Mahabad (also called *Sablax*¹), the historical and current sociopolitical center of the region, is the principal city of the dialect area. Other important towns in the speech zone of Mukri are Bokan, Sardasht (Kr. *Serdeşt*), Piranshahr (Kr. *Pîranşehr* and *Xanê*), Naqadeh (Kr. *Nexêde*), Shino or Oshnaviyeh (Kr. *Şino*), and Miandoab. The only extant statistics Lewis et al. (2013) put the number of Central Kurdish speakers in Iran as 3,250,000, approximately one third of which (i.e. more than one million) could be considered as speakers of Mukri. However, both of the estimates are meant to be provisional and should thus be treated with due caution.

The term “Mukri” is obtained from the name of the principality of *Mukriyan*, which ruled in the region from the late fourteenth to the late nineteenth century.² The names “Mukri” and “Mukriyani” denote provenance from the region, in addition to referring to the variety of Central Kurdish spoken there. We will refer to it as “Mukri” or “Mukri Kurdish”.

¹ This name comes from the name of the river *Sauj-boulagh* (or *Sauj-bulaq*), originally a Turkic name, which was replaced by Mahabad during the nationalist Pahlavi regime (Minorsky 1957:65), more precisely, in 1935 (Vali 2011:26).

² See Hassanpour (1989), but also Oberling (2010) and Minorsky (1957:73–74).

The speakers refer to their language as *kurdi* ‘Kurdish’, but they also employ names such as *mukri* and *mukriyāni* ‘Mukri Kurdish’ or a more learned form, *binzārāwey mukriy sorani* or *binzārāwey kurdiy navendi* ‘the Mukri sub-dialect of Sorani/Central Kurdish’.³

The Mukri variety has been observed to be “a very pure” form of (Central) Kurdish by orientalist such as Soane (1912:375, cited in Hassanpour 1992) or Nikitine (1956:164). It is the variety in which a number of influential authors wrote in the second half of the twentieth century, and served as the basis for the written language of Central Kurdish in Iran, used in publications and broadcasting. It is also claimed that Mukri has partially been the basis for the codification of modern standard Sorani Kurdish in Iraq and Iran (Hassanpour 1992: Ch. 8).

The variety is in close contact with Iranian Azeri in Naqadeh and with Kurmanji Kurdish in Oshnaviyeh. Mukri Kurdish is also spoken as a second language by portions of Iranian Azeri and Kurmanji-speaking communities, as well as by a small Kalhori-speaking⁴ community in the Mukriyan region.

Persian is the only official language and the medium of education in Iran, as laid down by Article 15 of the Constitution of the Islamic Republic of Iran. Although the same article allows for a restricted use of the regional and “tribal” languages in the educational system, in practice, they are rigidly excluded from such high domains (Sheyholislami 2012:31). In spite of the strict exclusion of Kurdish from education and other official restrictions on its use in Iran (see Hassanpour 1992; Sheyholislami 2012), the Mukri variety is a thriving medium of communication in its dialect zone in almost all of the social domains, excluding strictly official institutions such as courts of law. For instance, it is the sole language of the marketplace in predominantly Kurdish towns such as Mahabad (the center of the region), Oshnaviyeh, Bokan, and Piranshahr.

Mahabad and other towns of the Mukriyan region are administratively under Azerbaijan-e Gharbi (Western Azerbaijan) province, with its majority of Iranian-Azeri speakers. The de facto working language in the province is Iranian Azeri, and it is the main language which is promoted as part of the state-sanctioned cultural activities in local languages. The use of Kurdish in cultural activities is much more restricted than that of Iranian Azeri. There is thus one private language center where Kurdish grammar and literacy are taught, one state-sponsored bilingual (Kurdish-Persian) bi-monthly magazine (*Majallay Mahabad*) and one public radio station broadcasting a few hours a day in Mukri Kurdish (see also Sheyholislami 2012:32–37).

The speech community of Mukri is predominantly made up of bilinguals in Kurdish and Persian, whereas, especially in town-centers and areas with mixed populations, trilingualism in Kurdish, Persian and Iranian Azeri is widespread. In

³ For instance, in a recent book (Wetmani & Xelife 2009), describing the villages of Mahabad, written in Kurdish by two local researchers, the Mukri variety is specified as *binzaravey mukriy kurdiy sorani* ‘the Mukri sub-dialect of Sorani Kurdish’, where *sorani* ‘Sorani’ is the term covering the whole branch of CK and *mukri* ‘Mukri’ one of its dialects.

⁴ Kalhori is a dialect of Southern Kurdish. See for the speech zone of SK, and see Fattah (2000) for an extensive treatment of the grammar of SK dialects.

such areas, Persian, as the only official language in Iran, is mostly learned in school and by means of popular culture, whereas Azeri is functional in the marketplace. As a result of the dominant gender roles favoring males for higher social mobility, both forms of multilingualism seem to be more widespread among men than women. On the other hand, the speakers of Mukri and the dialects of Kurmanji,⁵ spoken in the northern lines of the dialect zone, are mostly able to communicate in their respective dialects usually without code-switching, but with some speech accommodation. If there is any code-switching at all, it is most often the speakers of Kurmanji who switch to Mukri Kurdish, which reflects the higher status of Mukri among the varieties of Kurdish.

1.3 The status of Mukri within Kurdish dialectology

The speech zone of Mukri corresponds to the north-east part of the Central Kurdish-speaking region, as shown in the map of Kurdish varieties in Figure 1. The status of Mukri as a distinct dialect of CK is acknowledged in most of the works on Kurdish,⁶ such as de Morgan (1904), Mann (1906), MacKenzie (1961), Hassanpour (1992), McCarus (2009), and Asatrian (2009), though the geographical extent of the dialect varies among the sources. For instance, MacKenzie (1961) divides CK dialects of Iraq into Suleimani, Warmawa, Bingird, Piždar, Arbil, Rewandiz, and Xošnaw, and adds Mukri as the CK dialect on the Iranian side which neighbors Rewandiz and Piždar dialects. He thus restricts it to Iran. With a different view, Hassanpour (1992) divides the CK speech area into two main dialect groups, namely, Mukriyani (our *Mukri*), comprising the northern half of CK, and Suleimani (“Suleimaniya” in the original), corresponding to the southern half of CK. This division implies the close affinity of Mukri with geographically neighboring varieties on the Iraqi side of the border (namely, the CK varieties spoken in Rewandiz and Piždar), a fact also emphasized in MacKenzie (1961).

McCarus (1958:4) states, in a note on Mann’s (1906) grammatical sketch of Mukri Kurdish, that Mukri is remarkably similar to Suleimani Kurdish. Hassanpour (1992:353–354) discusses main differences (i.e. isoglosses) between Suleimani (abbreviated Sul.) and Mukri dialects. Accordingly, in their phonologies, Mukri differs from Sul. in that the combination of [nd] is usually found as [ng] in the latter; the sound [d] in a large number of words and formatives in Mukri is realized as a centralized vocalic element [ə] in Sul. Note that both of the distinctions are related to the treatment of the /d/ phoneme in the language, which is generally unstable with different outputs in different dialects (see McCarus 2009:597 for discussion). In terms of morphology (cf. Hassanpour 1992:365–387; MacKenzie 1961; §2.3 below), to mention only a few distinctions, Sul. differs from Mukri in not having the oblique case suffixes in its nominal system; in the form of the indicative present and imperfective prefix (Mukri *de-* vs. Sul. *e-*); in its system of personal pronouns in that Mukri has an additional set of

⁵ Note that Kurmanji is referred to as *Šikāki* in the region, after a Kurmanji-speaking tribal confederation with the same name.

⁶ This ease at assigning a separate dialect status might also be due to the fact that the speech zone of Mukri corresponds to a socio-historically well-delimited region.

pronouns; and in its demonstratives, in that Sul. has distinctive forms of proximal and distal demonstratives, while Mukri only has a single demonstrative form.

In short, what is regarded as Mukri (or Mukriyani) may change depending on the level of variation taken into consideration. In this study, however, Mukri refers to the variety of CK spoken in the area extending from the south-east of Urmiya lake through Mahabad and Naqadeh in the center to Bokan and Sardasht in the east and south, and Piranshahr and Oshnaviyeh in the west. Thus, as Figure 1 indicates, the dialect zone is situated entirely in Iran.

The Mukri dialect has often been regarded as an outstanding one among all other Kurdish dialects, both by orientalists and native scholars. Thus, de Morgan (1904:2) claimed that Mukri is the best conserved variety of Kurdish in Iran on the grounds that it has been in contact with only Iranian Azeri, thus spared from the influence of other languages. E. B. Soane (1912) also considered that Mukri is very well “preserved” and should serve as the basis for the standard language (cited in Fossum 1919:8). Given the historical, cultural, and political importance of Mahabad for the Kurds of Iran, Mukri was also the accepted high code among the Kurds of the region in general, such that, as reported in Hassanpour (1992:163), it was used as the language of press even under the rule of the Kurmanji-speaking princes of the region at the turn of the nineteenth century.

Although modern standard Sorani or Central Kurdish is principally based on the Suleimani variety (cf. McCarus 1958:12–44 or Hassanpour 1992:351, Ch. 8), the phonological and morphological characteristics of the Mukri variety have also been largely integrated into the standard language (MacKenzie 1961:1), sometimes replacing the Suleimani forms.

It should be added to the above views on Mukri that, in terms of its lexical and grammatical features, it shows features that reflect its geographic proximity with Kurmanji dialects, indicating a degree of structural convergence, though less than the neighboring Soran/Rewandiz dialects on the Iraqi side of the frontier. In Mukri and the Soran/Rewandiz dialects, some of the case distinctions are preserved in the parts of the pronominal and nominal systems, bringing them closer to the case-based (more conservative) Kurmanji morphosyntax. In some varieties of Mukri, e.g. the variety spoken by the Dehbokri clan (Kr. *Dēbokrî*),⁷ the aspectual verbal particle is a preverb *we-*, similar to the Kurmanji preverb *ve-*, while it is a postverbal morpheme *-ewe* in the rest of CK and also in some neighboring Kurmanji dialects. As an areal feature of Mukri along with Soran/Rewandiz and neighboring Kurmanji varieties of Oshnaviyeh and Şemdinli (Kr. *Şemzînan*, cf. Haig & Öpengin forthc.), the *v-w* distinction is lost in favor of a generalized *w*, while the demonstrative system is different from that found in the rest of CK and NK. However, more comprehensive investigations are needed in order to state more about this presumed convergence zone between NK and CK. The Kurdish-Azeri contact in this region, where the bilingualism in the two languages is relatively common, especially in some towns, has not yet been studied, to the best

⁷ See Oberling (2010) and Hassanpour (1989) for some notes on this group, originally “a family of landed aristocrats”.

of our knowledge.⁸

As noted in previous literature, comprehension between speakers of Mukri and other varieties of CK is straightforward, while we are not informed about the questions of mutual intelligibility between the speakers of Mukri and Southern Kurdish as well as Gorani varieties. In the same vein, comprehension between NK and Mukri speakers is also a complex issue, since the linguistic repertoire of the speakers of these varieties is heavily influenced by the official language and the general context of the country in which the varieties are spoken. Accordingly, we assume that the NK speakers from Iran and Iraq have little difficulty, if any, in communicating with speakers of Mukri, while the NK speakers from Turkey and Syria would in general require a certain amount of exposure to the language before they can communicate with its speakers.

1.4 Previous work on Mukri and Central Kurdish

The Mukri variety has been the object of many studies, both by orientalist and native scholars.⁹ The first works relating to the Mukri dialect are Chodzko (1857), Houtum-Schindler (1884), de Morgan (1904), Mann (1906), and Fossum (1919). Chodzko (1857) is an impressionistic grammar sketch, while Houtum-Schindler's work consists mainly of a wordlist containing a mixed set of words from diverse Iranian languages. De Morgan (1904) is a dialectological survey of Kurdish varieties in Iran. Mann (1906:XXII–XXV) notes, however, that de Morgan's study is fraught with inconsistencies and outright errors, an opinion that we share as well (see Öpengin 2013:9–11).

Mann (1906), in contrast, is a comprehensive volume documenting principally the *bayts*¹⁰ of Mukri Kurdistan, with a few folktales, preceded by a brief sketch grammar of Mukri Kurdish in German. It is short, albeit precise, in treated issues, but the topics of grammar are unevenly discussed. Fossum (1919) is a grammar of Central Kurdish, with not much specific information on the dialect basis of the description. The work is not very useful, also for the additional reason that it invariably replicates the patterns and paradigms found in the study of classical languages such as Latin.¹¹

MacKenzie (1961) is a detailed comparative study of the dialects of Kurdish spoken in Iraq. Although MacKenzie himself did not extend his survey into the Mukri speech zone in Iran, he relied on Mann's (1906) data for his analysis of Mukri. His analysis is thus brief, and bears a number of inaccurate observations. Two other works specifically on Mukri are Eyubi and Smirnova (1968), in Russian, and Kelbasi (1983), which is mainly a collection of lexical items from the town

⁸ The following two words from our corpus are borrowings from Iranian Azeri: *yermâsi* 'potato' < *yerilmâsi* (lit. 'earth apple'); *qewâtûn* 'a late breakfast' < *gehwalti*.

⁹ See Öpengin (2013:9-14) for an evaluation of the previous work on Mukri and CK.

¹⁰ In Hassanpour's (1990) definition, "[a] *bayt* is an orally transmitted story which is either entirely sung or is a combination of sung verse and spoken prose." It is probably the most widespread and appreciated form of oral literature in the region.

¹¹ L. O. Fossum also translated, with the help of his local associates, a few books of the Bible into Mukri Kurdish in 1914, published in 1919 (cf. Thomas 1989).

center of Mahabad with some preliminary notes on phonology and grammar. Blau (2000) is a manual for learners of Sorani, while Thackston (2006) is a practical reference grammar of standard Suleimani Kurdish. Both of these works primarily address the needs of learners of the language. Haçî Marif (2001) is a highly detailed grammar of CK with many aspects of comparison to other Kurdish varieties. It is in Sorani Kurdish and contains many observations that would easily escape an outsider's attention. Finally, McCarus (2009) is a concise and fairly complete grammatical sketch of Suleimani Kurdish, which is basically very close to the standard use of the language in the media.

1.5 Fieldwork

The language material used in the present study was collected during three fieldwork trips to the Mukriyan region in Iran, which I undertook within the frame of my doctoral research from 2009 to 2013 (see Öpengin 2013). The first field trip took place in March-April 2011 (thirty-five days),¹² the second in October-November 2011 (twenty-five days), and the third in September 2012 (ten days). Between fieldwork trips, I retained contact with a number of native speakers via the Internet and was able to check many details of the analysis in this manner. The fieldwork was conducted principally in two villages of Mahabad, Sarewanan (twelve kilometers east of Mahabad) and Qozluje (thirty-two kilometers south of Mahabad), and in Mahabad and Oshnaviyeh town centers. These fieldwork localities are shown in Figure 2.

As a native speaker of the NK dialect of Şemdinli, close to the Iranian-Turkish border, and with a working knowledge of standard Sorani, I had a solid foundation for acquiring Mukri as well as considerable familiarity with the cultural context prior to beginning fieldwork there. After two weeks of work with two young native speakers, mainly to improve my knowledge of Mukri, I was introduced to a number of local researchers, one of them being Mr. Salah Payanyani, who had already conducted extensive fieldwork in the region for his project of a dictionary of Mukri oral literature (see Payanyani 2006, 2008, 2011). Besides graciously hosting me in the Mahabad town center, sharing his extensive knowledge on his language and the oral literature of the region, he also kindly introduced me to some of his informants in the villages around Mahabad. The two villages where I collected the main bulk of connected speech data were Sarewanan and Qozluje. To Qozluje, I took only day trips, while I spent a good part of my fieldwork time in Sarewanan, where I was generously hosted by Mr. Mihemed Ehmediazar, a self-taught poet and prose writer in his native tongue (see Ch. 4).

¹² See the report on the following link for a description of this first field trip in the region: http://lacito.vjf.cnrs.fr/INTRANET/documentspdf/Opengin_mission11.pdf [October 20, 2013]

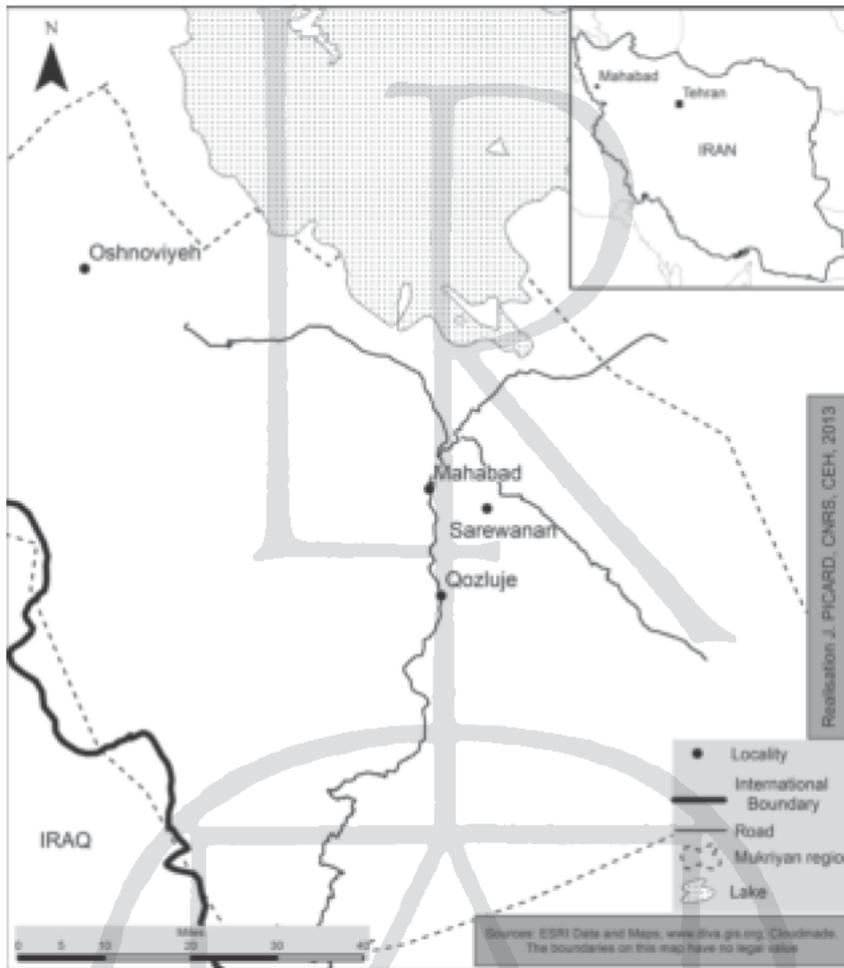


Figure 2. Fieldwork localities in the Mukriyan region of northwestern Iran

In both of the villages, the residents are mainly farmers. They cultivate grains, lentils, and vegetables, both for their own consumption and for selling in the market. All but a few families in the villages possess some plots of land. Some of them have recently started to convert their lands into orchards. Most of the residents also have some livestock such as sheep and goats, and some have cows, mainly for obtaining dairy products but also for professional stockbreeding. A common practice among those who have livestock is to go to the highlands (*hewār*) from mid-May to mid-September, while the shepherds also spend October in the highlands. Sarewanan is well connected to the town center, such that most of the villagers have some relatives in the town center and frequently go there for shopping, medical care, school, and similar purposes.

1.6 Corpus

The corpus used in this grammatical description is part of the larger spoken Mukri Kurdish database, described in detail in Öpengin (2013). The corpus used here is

made up of seventeen narrative texts, of varying length (from shorter than two minutes to about forty minutes). They make up a total of 165 minutes of running texts, comprising about 23,576 word tokens. The ensemble of the texts is presented in Table 1, while further details on the storyline and other dimensions of the texts are provided in Chapter 4.

Table 1. The corpus of spoken texts in Mukri Kurdish

Text code	Database extension	Length (mn)	Genre	Narrator	
				Name	M-F/Age
HF	Qad 3 HF	04:01	folktale	Qadir	M/73
QM	Qad 7 QM	02:36	procedural text	Qadir	M/73
QA	Qad 9 QA	07:50	autobiography	Qadir	M/73
EP	Qad 11 EP	39:42	folktale	Qadir	M/73
NZ	Qad 13 NZ	10:43	folktale	Qadir	M/73
ČQ	Qad 20 ČQ	16:46	folktale	Qadir	M/73
JA	Qad 17 JA	03:50	folktale	Qadir	M/73
MK	Os 2 MK	16:44	folktale	Osman	M/62
MN	Os 3 FM	08:18	folktale	Osman	M/62
KF	Mih 3 KF	06:16	folktale	Mihemed	M/35
ŽB	Mih 4 ŽB	10:54	real-life story	Mihemed	M/35
TS	Mih 6 TS	01:49	procedural text	Mihemed	M/35
ČN	Mih 8 ČN	16:23	folktale	Mihemed	M/35
GD	Mih 10 GD	05:18	anecdote	Mihemed	M/35
TA	Mih 9 TA	05:07	anecdote	Mihemed	M/35
KR	Hej 1 KR	02:41	anecdote	Hejar	M/9
HX	Hel 1 HX	05:43	autobiography	Helim	F/65
Total	17 texts	165 mns	5 genres	5 narrators	Range 9 – 73

The texts of the corpus were collected mainly from three speakers. First, Qadir Mihemmed Resuli is a seventy-three-year-old male farmer, who moved to Sarewanan from another village in the region some forty years ago. He is mostly monolingual in Kurdish, with a little knowledge of Persian. He has spent all his life in the region. Second, Osman Remezani is a sixty-two-year-old resident of Sarewanan. He is a farmer and lives between Mahabad town center and Sarewanan village. In addition to Kurdish, he also speaks Persian. Third, Mihemed Ehmediazar is a thirty-five-year-old male. He is a farmer and stockbreeder. He has studied up to the high school level and is bilingual in Mukri Kurdish and Persian, and also fully literate in standard Sorani Kurdish. In addition to these sources, one

text was collected from a sixty-five-year-old woman and another text from a nine-year-old male child. The texts are of various genres: folktales, autobiographical narratives, procedure texts, (regional) anecdotes, real-life stories, etc. The data was collected mainly by the researcher, but in some of the interviews, especially at the initial stages of the fieldwork, my hosts Mr. Salah Payanyani and Mr. Mihemmed Ehmediazar were also present in addition to the informants. There are also a high number of notes that I took in the field during daily interactions with speakers. These notes are not included in the main corpus of the study, but I have occasionally made use of them in minor ways, showing them in the text indicated with the abbreviation FN [Field Note].

The data consists of audio recordings in WAV (Waveform Audio File) format. I used a Zoom h2 Handy Recorder as the principal recording device, while I had an additional audio recorder switched on during the sessions, recording in MP3 format. The collected data was systematically saved on a laptop, together with notes on informants and the settings, and backed-up on two external hard drives. The audio recordings for the ten texts of the corpus (in Table 1) that are presented in Ch. 4 are provided on the CD that accompanies this book.¹³

Processing the corpus data involved a number of stages. First, these seventeen texts were selected out of the larger Mukri Kurdish database for analysis and publication on the basis of linguistic interest of the text (e.g. presence of various types of person reference, dialogues), theme, and quality of the recording. I transcribed the initial two texts with the help of my informants and consulted with them to check the accuracy of my transcription for most of the texts in the corpus. I then provided an English translation for the texts. The transcription and translation were done in the software program of ELAN (EUDICO Linguistic Annotator), with an orthography based on Latin script, which is an adapted version of the widespread orthography in use for the study of Iranian languages (e.g. Mahmoudveysi et al. 2012). The aligned transcription and translations were then transferred into Toolbox for glossing and running diverse grammatical and lexical analyses. Each annotation unit (which roughly, though not necessarily, corresponds to a sentence) has an identification code in the corpus. It is composed of the text code and the number of the annotation unit in the given text. For instance, the code MK.40 is to be interpreted as the sentence/annotation unit no. 40 of the text MK (to be found in the text code column in Table 1, see also Ch. 4). The identification codes figure on the right-most edge of the first line of each example sentence in the grammar. This allows locating the given example in its context, in both the transcription and audio versions.

¹³ One of the texts, text code HF in Table 1, is also available on the PANGLOSS archive of LACITO, in the open access category, available at: lacito.vjf.cnrs.fr/archivage/languages/Kurdish.htm [October 20, 2013]