

Language Specific Peculiarities Document for Kurmanji Kurdish as Spoken in Turkey

1. Special handling of dialects

Kurmanji Kurdish is a major branch of modern Kurdish, which belongs to the Iranian group of languages. Kurdish is largely a spoken language with a limited but growing body of modern literature. There are many dialectal varieties of Kurdish spread over a wide area of Turkey, Iraq, Syria, Iran, Armenia and Azerbaijan. There are, in addition, different classifications of Kurdish languages and Kurdish peoples that align dialectal choice with tribal affinity. Linguistic and Kurdish scholars usually describe modern Kurdish as having two closely related major branches: Kurmanji (the northern branch) and Sorani (the southern branch). They are not mutually intelligible (see Thackston (2006) p. vii). The Kurdish Institute summarizes the situation as follows: “Kurdish has two regional standards, namely Kurmanji in Turkey, and Sorani farther east and south. Roughly half of Kurdish speakers live in Turkey.” (see <http://www.institutkurde.org/en/> and also <http://www.blueglobetranslations.com/about-kurdish-kurmanji-language.html>). Within the larger region there are two other languages often associated with ethnic Kurds. These are Dimili (also known as Zazaki) and Gorani. Although sometimes classified as sub-dialects of Kurdish (e.g., Kurdish Language - Britannica Online Encyclopedia), these languages belong to a different group of Iranian languages (see for example, <http://kurds.history.enacademic.com/346/Kurmanji>).

Although Kurmanji Kurdish is spoken across a range of countries, it has the advantage of being spoken by around 60-80% of Kurds and is considered a regional standard. The majority of Kurmanji speakers live in Turkey, making it the ideal country for collection of audio data (see for example, <http://linguakurd.blogfa.com/post-106.aspx>).

This collection took place in Turkey. Within Turkey there are dozens of spoken varieties of Kurmanji and dialect choice for this collection was dependent not just on mutual intelligibility and population density, but also on the influence and spread of standardization initiatives for the written forms of the language within Turkey. There is no single standard dialect of Kurmanji, although Kurmanji speakers often identify with a standard known as Botanî Kurmanji. This variety has attained a certain status as a modern standard due to increasing use in broadcasting through various Kurdish radio and TV channels, primarily in eastern Turkey (see Haig (2006), p. 283). Botanî Kurmanji is also considered to be very similar to the Hakkari dialect of Kurmanji. Middle Kurmanji is mutually intelligible with these other Kurmanji dialects. Although it does add some vocabulary variation to the collection (spoken in Diyarbakir and western regions of Mardin and Urfa), it is spoken in relatively populous areas, thus aiding in data collection. The inclusion of Western Kurmanji (spoken in Adiyaman, Sanlurfa and Gaziantep) would have diluted the Kurmanji standard somewhat because of the extensive use of non-Kurmanji words, particularly Turkish, in everyday spoken language (Personal

communication with language consultants).

In order to create a collection of reasonably standard spoken Kurmanji (given the population density of Kurdish speakers within Turkey), the collection focused on speakers from the so-called Southeastern and Eastern Anatolian region of Turkey. This region includes speakers from the provinces Sirnak, Siirt, Mardin, Hakkari, Van and some regions of Batman and Bitlis. Middle Kurmanji includes speakers from Diyarbakir and surrounds. In order to avoid Western Kurmanji, the collection excluded speakers from Andiyaman, and Gaziantap. Please note that Dimili (also known as Zazaki) and Gorani were also excluded from the collection because of low levels of mutual intelligibility with Kurmanji.

Kurmanji Dialect Name	Turkish Provinces Where Spoken
Botanî	Cizre in Sirnak, Siirt region, eastern Mardin, south-east of Diyarbakir, parts of Batman city, southern and eastern parts of Bitlis and in southern Van Province.
Hakkâri	Hakkâri Province
Middle Kurmanji	Greater Diyarbakir and western areas of Mardin and Urfa (Şanlıurfa).

2. Deviation from native-speaker principle

Only native Kurmanji Kurdish speakers from the south-eastern and eastern Anatolian region of Turkey were collected in this project.

3. Special handling of spelling

Because Kurmanji is predominantly a spoken language, there were spelling standardization issues to resolve. Foreign words (including, but not limited to, Turkish, Arabic, Persian and English) have been standardized to Kurdish spelling. These words are transcribed in this database due to the large amount of code-switching with Turkish forms. In preparing this database, we have used the services of an experienced language consultant who is highly familiar with existing standardization proposals for Kurmanji. In cases where a single accepted form does not exist, we have made a choice between forms in as principled a way as possible, while taking into account such factors as frequency of use and native speaker preference.

4. Description of character set used for orthographic transcription

Kurmanji has been, and is written, in a variety of alphabets. The most common as of today is Kurmanji used in Turkey and Europe, which is written in a modified Turkish Latin alphabet (see Thackston (2006), p. ix). There are two Latin alphabets that could be used for Kurmanji according to the Kurdish Academy of Language. One is a Latin alphabet designed for all varieties of Kurdish, known as the Unified script or Yekgirtú; however, this has yet to gain universal usage for Kurmanji. The Latin alphabet Appen used is the North Kurdish Latin script,

also known as Latin Kurmanjî (see <http://www.kurdishacademy.org/?q=node/145> and also <http://www.omniglot.com/writing/kurdish.htm>). This script is most commonly used for written Kurmanji in Turkey and Syria, as well as in exile communities in Europe and the United States of America, Canada, and Australia. Most early modern Kurdish literature uses this script (see <http://linguakurd.blogfa.com/post-106.aspx>).

North Kurdish Latin Script	North Kurdish Latin Script
A a	N n
B b	O o
Ç ç	P p
D d	Q q
E e	R r
Ê ê	S s
F f	Ş ş
G g	T t
H h	U u
I i	Û û
Î î	V v
C c	W w
J j	X x
K k	Y y
L l	Z z
M m	

Latin Kurmanjî uses the Basic Latin unicode range (U+0041-U+007A) plus a number of forms from the Latin Extended_A set. The Latin Extended_A characters are included in the table below.

Unicode	Character	Unicode	Character
U+00C7	Ç	U+00E7	ç
U+00CA	Ê	U+00EA	ê
U+00CE	Î	U+00EE	î
U+00DB	Û	U+00FB	û
U+015E	Ş	U+015F	ş

5. Description of Romanization scheme

Not applicable.

6. Description of method for word boundary detection

Word boundaries in the orthography are determined by location of white spaces.

7. Table containing all phonemes in the stipulated notation

The phonemic transcription of the words in this database uses X-SAMPA symbols (see <http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm>). There are thirty-one phonemes in Kurmanji (plus up to eleven allophones and foreign phones). The lexicon uses twenty-three consonant phonemes including two semi-vowels, plus four consonant allophones and foreign phones. It also uses eight vowel phonemes plus two additional vowel allophones, all monophthongs.

KURMANJI KURDISH PHONE CHART

KURMANJI CHARACTER	IPA	X-SAMPA	Example
CONSONANTS			
p	p	p	pey
b	b	b	ber
t	t	t	ta
d	d	d	di
k	k	k	kes
g	g	g	gel
q	q	q	qey
f	f	f	fen
v	v	v	vir
s	s	s	sar
z	z	z	zar
ş	ʃ	S	şal
j	ʒ	Z	jêr
x	x	x	xwe
h	h	h	hay
ç	tʃ	tS	çar
c	dʒ	dZ	car
m	m	m	maf
n	n	n	nan

KURMANJI CHARACTER	IPA	X-SAMPA	Example
w	w	w	wek
y	j	j	yek
l	l	l	lez
r	r	4	ber
CONSONANT ALLOPHONES			
r	r	r	roj
FOREIGN PHONES			
empty string	ɸ	?\	cemaet
x	ɣ	G	xem
h	ħ	X\	hakim
VOWELS			
î	i	i	nîv
i	ɪ	l	dil
ê	e	e	hêz
e	æ	{	bez
û	u	u	bûn
u	ʊ	U	du
o	o	o	nod
a	a	A	bav
VOWEL ALLOPHONES			
i	ɨ	1	vedixwim
e	ɛ	E	odeya

OTHER SYMBOLS	
“	primary stress
.	syllable break
#	word boundary

Consonant Allophones in Kurmanji

Although some resources (e.g., Thackston (2006), pp. 2-4) posit a phonemic contrast between aspirated and pharyngealized variants of /p/, /t/, /k/, and /tʃ/, such distinctions did not occur extensively in our database, and hence are not made in either our orthographic transcriptions or our pronunciation lexicon.

There is a trill allophone [r] of the flap/tap /r/ that occurs in word-initial position and sometimes in other positions. This allophone has been represented in the lexicon mainly in word-initial position.

Foreign Consonant Phones in Kurmanji

Although dialect selection has been designed to minimize the influence of other languages on Kurmanji, pronunciation and vocabulary can be influenced by knowledge and usage of Turkish and Arabic. The set of three foreign consonant phones listed in the table above generally arise from Arabic borrowings. The foreign phone [ɣ] is also found in Turkish and may arise from Turkish borrowings in the data.

/ɣ/

The Arabic phone (known as ‘ayn – ع), a voiced pharyngeal fricative [ɣ], is sometimes used in the pronunciation of words borrowed from Arabic. This phone has been represented in the lexicon as [ʔ\].

/ɣ/

The voiced post-velar fricative [ɣ] (from Arabic and Turkish, known in Arabic as ghayn – غ) is also attested to occur in the pronunciation of borrowed words. This phone has been represented in the lexicon as [G].

/ħ/

Some speakers of Kurmanji pronounce the voiceless pharyngeal fricative [ħ] (Arabic ح) found in some words borrowed from Arabic. This phone has been represented in the lexicon as [X\], always as a pronunciation variant of the native Kurdish (glottal) fricative [h].

Vowel Allophones in Kurmanji

[i]

Given the close proximity to spoken Turkish, some speakers use a close central unrounded vowel [i] similar to the Turkish character ‘ı’. This allophone is usually attested in Kurmanji dialects outside of the collection area for this project (Ozsoy, A. Sumru, and Yektan Tiirkyilmaz, 2006).¹ This vowel seems to occur in limited phonological environments, for example, where ‘ı’ occurs in closed syllables ending in /m/.² This variation occurs in the recorded data and has

¹ See Ozsoy et al. (2006), p. 300.

² Vowel allophones are modeled on Thackston (2006), p. 1.

been included as a variant pronunciation in the lexicon represented as [1].

[ə], [ɛ]

Although the Kurmanji character ‘e’ is usually pronounced as /æ/, when it occurs in the spelled sequence 'ew' it is sometimes pronounced as [ə]. This allophone is not salient and is not represented in the lexicon as a variation. (The representation [æw] – X-SAMPA [w] -- was used instead.) When ‘e’ is followed by ‘y’ but not in the same syllable, it can be pronounced as [ɛ]; and in post-stress positions, particularly word-finally, it can also be pronounced as [ɛ]. This variation is represented in the lexicon.

8. Other language specific items

It is common in predominantly spoken languages like Kurmanji for there to be extensive lexical variation between so-called mutually intelligible dialects. In fact, in this region, speakers pride themselves in knowing a range of lexical “synonyms” that help identify the origin of a particular speaker or his affiliation with a particular place.

As will be evident from the tables below there is often one or more alternative spellings of a single form in common usage. Additional lexicalizations may also be commonly used to confer the same meaning. In the latter case we have added alternative lexemes or phrasal equivalents that may occur in the collection (Common variations of the form), particularly in the spontaneous speech component of the database. We have chosen to represent these with only one standardized form, though variant spellings occur for many of these lexemes. For example, ‘deh’ meaning the number 10 may be spelled with or without the final ‘h’; we have standardized to the form with ‘h’. Similarly, a number of words containing the letter sequence ‘zd’ may also be spelled with ‘nzd’; we have standardized to ‘zd’.

We have also noted in the table below additional standardization decisions taken during preparation of the final database. The standardized spellings are indicated in brackets beside the items in question. The principle generally applied in these cases is that words with the same meaning that differ by one letter/sound only will be standardized to a single form. In a few cases (e.g., yek/êk), although the words differ by more than one sound, they were judged similar enough to be conflated to a single spelling, and the two variants are given as alternate pronunciations in the pronunciation dictionary.

8.1 Table of Digits

Digit	Kurmanji (Appen variant)	Common variations of the form
0	sifir	
1	yek	êk (alternate pronunciation)
2	du	didu
3	sê	sisê
4	çar	
5	pênc	
6	şeş	
7	heft	
8	heşt	
9	neh	

8.2 Other Numbers

Digit	Kurmanji (Appen variant)	Common variations of the form
10	deh	
11	yazdeh	
12	duwazdeh	
13	sêzdeh	
14	çardeh	
15	pazdeh	
16	şazdeh	
17	hivdeh	
18	hijdeh	
19	nozdeh	
20	bîst	
30	sih	
40	çil	
50	pêncî	
60	şêst	
70	heftê	
80	heştê	
90	nod	nehwêd
100	sed	
1,000	hezar	
10,000	deh hezar	
100,000	sed hezar	
1,000,000	milyon	milyonek

9. References

- Chyet, M. L. (2003). *Kurdish-English Dictionary (Ferhengê Kurmancî – Ingîzî)*, New Haven: Yale University Press.
- Ekici, D. (2007). *Kurmanji Kurdish Reader*, Hyattsville, Maryland, USA: Dunwoody Press.
- Haig, G. (2006). "Turkish Influence on Kurmanji: Evidence from the Tunceli Dialect." *Turkic-Iranian Contact Areas: Historical and Linguistic Aspects*, 62.
- MacKenzie, D. N. (1986). "Language in Kurds & Kurdistan", *Encyclopaedia of Islam*, Vol V.
- Ozsoy, A. Sumru, and Yektan Tiirkyilmaz, (2006). "Front rounded vowels in the Sinemili dialect of Kurmanji-a case of language contact?" *Turkic-Iranian Contact Areas: Historical and Linguistic Aspects*, 62.
- Saadallah, S. (2000). *Saladin's English-Kurdish Dictionary*. Istanbul: Avesta Press.
- Soane, E.B. (1913) *Grammar of the Kurmanji or Kurdish Language*, Part I, p 5, London.
- Thackston, W. M. (2006). *Kurmanji Kurdish: A Reference Grammar with Selected Readings*. (Harvard University, Department of Near Eastern Languages).

Online Resources

<http://www.blueglobetranslations.com/about-kurdish-kurmanji-language.html>

<http://www.institutkurde.org/en/>

<http://www.kurdishacademy.org/?q=node/145>

http://kurds_history.enacademic.com/346/Kurmanji

<http://linguakurd.blogfa.com/post-106.aspx>

<http://www.mideastweb.org/Middle-East-Encyclopedia/kurds.htm>

<http://www.omniglot.com/writing/kurdish.htm>

<http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm>