Multi-CAST

Multilingual Corpus of Annotated Spoken Texts

Northern Kurdish

— metadata sheet —

June 2016

Geoffrey Haig
University of Bamberg

Hanna Thiele
Bamberg/Freiburg

cite this document as:

Citing the corpus

The Multi-CAST Northern Kurdish corpus should be cited as follows:

```
Haig, Geoffrey & Thiele, Hanna. 2016. Northern Kurdish.
  In Haig, Geoffrey & Schnell, Stefan (eds.),
  Multi-CAST (Multilingual Corpus of Annotated Spoken Texts).
  (https://lac.uni-koeln.de/en/multicast-nkurd/)
  (date accessed.)
```

Language data

language	English (ISO 639-3: kmr)
affiliation	Indo-European, Iranian, Northwestern
area spoken	Turkey, East / Iraq, North / Iran, West
varieties recorded	Northern Kurmanji, Erzurum / Muş
text type	traditional narratives
sources	Haig & Öpengin, forthcoming; Öpengin & Haig 2014

Corpus data

file name	speaker	date rec'd	text type	length h:mm:ss	clause units
nkurd_muserz01	NK01	2000	traditional	19:53	619
nkurd_muserz02	NK02	2002	traditional	12:12	482
		totals for the Northern Kurdish corpus		32:05	1,101

Table 1. List of files in the Northern Kurdish corpus.

speaker	gender	age	date born	date rec'd	clause units	texts
NK01	male	~50	~1950	2000	619	*nkurd_muserz01*
NK02	female	~60	~1940	2002	482	*nkurd_muserz02*

Table 2. List of Northern Kurdish corpus speakers.

Background to the recordings

nkurd_muserz01

This text was recorded by Geoffrey Haig (GH) with a speaker called Miheme, who grew up in a village near Muş. The speaker had left Turkey approximately

10 years previously and had since settled in Germany; the recordings were made in his allotment garden in Kiel, North Germany, in the company of Miheme's wife and another friend of the family. GH made a long series of recordings with Miheme, most of which have been transcribed and translated by GH with the assistance of native speakers.

The stories are Miheme's renderings of traditional Kurdish folkloric texts. Although not a trained story-teller, Miheme relished the opportunity to tell these stories, most of which he was recalling from childhood memories. He had no qualms about embellishing them in various ways when his memory failed him. His Kurdish is quite strongly in influenced by Turkish, his main language of communication over the past two decades, but he is undoubtedly a fluent speaker of Kurmanji.

nkurd_muserz02

This text was recorded by Abdullah Incekan in 2002 in Essen, Germany; the speaker is his grandmother Güllü Tunç, who was visiting Germany at the time. The atmosphere was relaxed; a number of family members including small children were present during the recordings. The speaker is a monolingual Kurmanji speaker who has spent her lifetime in a village of the region Tekman, south of Erzurum. The text was transcribed by Abdullah Incekan and GH, and translated by GH.

Content-wise, this text is obviously related to the well-known fairy tale *Cinderella*, and contains key motifs such as the evil stepmother, the slipper, the prince and so on, but the latter part of the story seems to stem from a different source, and at times the narrative lacks coherence.

References

Haig, Geoffrey & Öpengin, Ergin. forthcoming. Kurmanji Kurdish in Turkey: structure, variation and status. In Bulut, Christiane (ed.), *The minority languages in Turkey*. Wiesbaden: Harrassowitz.

Öpengin, Ergin & Haig, Geoffrey. 2014. Regional variation in Kurmanji: A preliminary classification of dialects. *Kurdish Studies* 2(2). 143–176.

Corpus changelog

date	changes	affected texts
2015–05	added 2 corpus files	nkurd_*
		muserz01, muserz02