

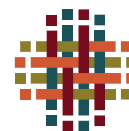
Multi-CAST

Northern Kurdish annotation notes

Geoffrey Haig
University of Bamberg

May 2019
v1.1

Hanna Thiele
University of Bamberg / University of Freiburg



ARC CENTRE OF EXCELLENCE FOR
THE DYNAMICS OF LANGUAGE



Australian Government
Australian Research Council



University of Bamberg

DFG

Multi-CAST

Multilingual Corpus of
Annotated Spoken Texts

Citation for this document

Haig, Geoffrey & Thiele, Hanna. 2019. Multi-CAST Northern Kurdish annotation notes. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

Citation for the Multi-CAST collection

Haig, Geoffrey & Schnell, Stefan (eds.). 2015. *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.

The entirety of Multi-CAST, including this document, is published under the *Creative Commons Attribution 4.0 International Licence* (CC BY 4.0), unless noted otherwise. The licensing terms can be reviewed online at creativecommons.org/licenses/by/4.0/.

Multi-CAST Northern Kurdish annotation notes v1.1 last updated 8 May 2019
This document was typeset by NNS with X_YL^AT_EX and the *multicast3* class (v3.0.9045).

Contents

1	Notes on the GRAID annotations	1
1.1	Transitive and intransitive verbs: Verbs of speech	1
1.2	Light verb constructions	1
1.3	Adpositions	1
1.4	The verb <i>anîn</i> ‘bring’	2
	References	2
	Appendices	3
A	List of corpus-specific GRAID symbols	3
B	List of abbreviated morphological glosses	4

1 Notes on the GRAID annotations

This document describes the implementation of the GRAID annotation conventions (Haig & Schnell 2014) in the Multi-CAST Northern Kurdish corpus. It corresponds to version 1905 of the annotations, published in May 2019. Unless a more recent version of this document exists, it also applies to any later versions of the annotations.

1.1 Transitive and intransitive verbs: Verbs of speech

Transitive and intransitive verbs can be distinguished on morphological grounds: transitive verbs trigger the ergative construction in the past tenses, while intransitives do not. In general, this was taken as the main diagnostic for distinguishing transitive and intransitive predicates, with one main exception: verbs of speech, in particular *gotin* ‘say’. Morphologically, it definitely belongs in the class of transitive verbs. However, it is very rarely used with a nominal object; rather, it generally introduces direct speech, which may continue for several clauses (and hence is difficult to classify as an “object”). Given the massive frequency of this verb, we have decided to treat it as intransitive, but have glossed all forms with $\langle ds_v : pred \rangle$, indicating that the verb concerned is a verb of speech. Likewise, the subjects of this verb have been glossed as $\langle : s_ds \rangle$, indicating that they are treated as a special case of intransitive subjects (S). When interpreting the data, analysts may choose to include these subjects in the S-category, or in the A-category, or treat them as a special case, or ignore them entirely.

- (1) *qîzikê* *gotîye:* *perê* *min* *pir* *e*
 girl said money.of mine much is
 # np.h:s_ds ds_v:pred #ds np:s rn_pro.1:poss other:pred cop
 ‘The girl said, “I have plenty of money...”’ [mc_nkurd_muserz01_0115]

1.2 Light verb constructions

Northern Kurdish makes extensive use of light verb constructions: some non-verbal element is combined with one of a small number of so-called light verbs, most commonly ‘do, be, give, strike, bring,’ to create a new predicate. Many of these can conveniently be treated as simple predicates, and in some cases they are orthographically treated as one word. In other cases the borderline to other combinations of generic object plus verb is difficult to draw. For highly frequent combinations, particularly if the non-verbal element does not occur outside of the particular combination, we have glossed the non-verb combination with $\langle : lvc \rangle$:

- (2) *em* *ê* *ji* *xwe* *ra* *xwedî* *kin*
 we will for ourselves POP raising 0_her do
 #ds pro.1:a aux adp refl_pro.1:obl pop_adp np:lvc 0.h:p v:pred
 “We will raise (her) for ourselves.” [mc_nkurd_muserz01_0069]

1.3 Adpositions

Northern Kurdish has quite a number of circumpositions consisting of a preposition plus a postpositional clitic particle. The former is glossed $\langle adp \rangle$, the latter $\langle pop_adp \rangle$; see (2) above for an example.

1.4 The verb *anîn* ‘bring’

In the dialect of Northern Kurdish recorded for this corpus, the verb *anîn* ‘bring’ is commonly used in combination with another verb, most often *come*. In order to express *bring, fetch*, etc., we find *tîne tê* ‘(lit.) brings comes’. To avoid inflating the number of clauses, we have treated this combination as a lexicalized one, thus analyzing such expressions as a single clause, and have glossed the form of *anîn* with ⟨*vother*⟩.

- (3) *baş e* *tînin tîn* *hundurê malê*
 0_it good is 0_the 0_it bringing come to.inside.of
 #ds 0:s other:pred cop # 0.h:a 0:p *vother* *v:pred* adp
 ‘(She said) “It’s fine”; (they) bring (it) into the house.’ [mc_nkurd_muserz01_0061]

References

- Haig, Geoffrey & Öpengin, Ergin. 2018. Kurmanji in Turkey: Structure, varieties, and status. In Bulut, Christiane (ed.), *Linguistic minorities in Turkey and Turkic-speaking minorities of the peripheries*, vol. 111. Wiesbaden: Harrassowitz.
- Haig, Geoffrey & Schnell, Stefan. 2014. *Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators (version 7.0)*. (<http://multicast.aspra.uni-bamberg.de/#annotations>) (Accessed 2019-03-08).
- Haig, Geoffrey & Schnell, Stefan (eds.). 2016. *Multi-CAST: Multilingual Corpus of Annotated Spoken Texts*. (<http://multicast.aspra.uni-bamberg.de/>) (Accessed 2019-03-08).
- Haig, Geoffrey & Thiele, Hanna. 2015. Multi-CAST Northern Kurdish. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (<http://multicast.aspra.uni-bamberg.de/#nkurd>) (Accessed 2019-03-08).
- Öpengin, Ergin & Haig, Geoffrey. 2014. Regional variation in Kurmanji: A preliminary classification of dialects. *Kurdish Studies* 2(2). 143–176.

Appendices

A List of corpus-specific GRAID symbols

The following is a list of the non-standard GRAID symbols used in the annotation of the Multi-CAST Northern Kurdish corpus. Please refer to the *GRAID manual* (Haig & Schnell 2014: 54–55) for an inventory of basic GRAID symbols.

Form symbols and specifiers

<intrg_pro>	interrogative pronoun
<pn_np>	proper name
<ds_v>	verb of speaking

Function symbols and specifiers

<:a_ds>	subject of a transitive verb of speech
<:s_ds>	subject of a intransitive verb of speech

Clause boundary symbols

<rep>	clause unit that is a verbatim repetition of a preceding clause unit
-------	--

Subconstituent symbols

<ln_dem>	demonstrative pronoun
----------	-----------------------

Other symbols

<excl>	exclamations
<redup1>	second part of a partial reduplication; stylistic figure used for emphasis etc.
<pop_adp>	postpositional particle; co-occurs with a preposition (see (2) above)

B List of abbreviated morphological glosses

MODAL	modal	RDP	partial reduplication
PL	plural		
POP	postpositional particle	NC	not classified

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/

