An Empirical Linguistic Analysis of Clitics in Kurdish

Sina Ahmadi, Antonios Anastasopoulos, Géraldine Walther

{sahmad46,antonis,gwalthe}@gmu.edu

Highlights

- **Task: Analyze the complexities of Central Kurdish endoclitic markers**
- **Two types of endoclitics:**
  - Pronominal markers
  - Emphatic clitic =*îş*
- **Contributions:**
  - Document rare phenomena
  - Tagging a corpus for Sorani
  - Data-driven analysis
  - Cross-lingual comparison

https://github.com/sinaahmadi/EmpiricalClitics

Clitics in Kurdish

Kurdish is known for displaying a number of morphological complexities related to clitics, particularly **endoclitics**. The placement of these clitics is dependent on:

*complex subject vs. object person/number properties and tense & **split-ergativity***

0					girt					past stem of GIRTIN (to take, to get)
1					girt	im				I got
2					girt	im	in			I got them
3					girt	im	in	e		I got them to/with
4					girt	im	in	e	ewe	I got them to/with again
5			ne	im	girt	in	e	ewe		I did not get them to/with again
6		ne	im	de	girt	in	e	ewe		I was not getting them to/with again
7	heł	im	ne	de	girt	in	e	ewe		I was not lifting them to/with again

the complex patterns of pronominal endoclitics

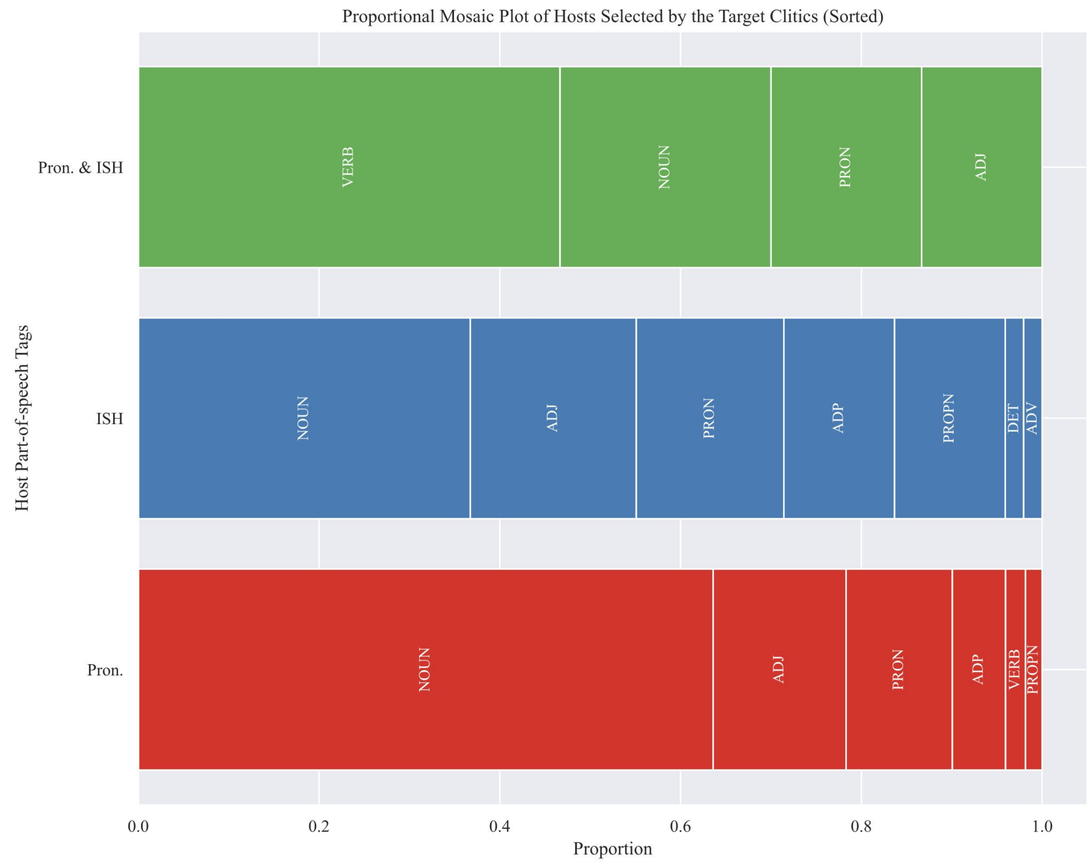
0						girt					past stem of GIRTIN (to take, to get)
1						girt	im				I got
2						girt	im	in			I got them
3						girt	im	in	e		I got them to/with
4						girt	im	in	e	ewe	I got them to/with again
5						girt	îş	im	in	e	ewe	I got them also to/with again
6			ne	îş	im	girt	in	e	ewe		I did not get them also to/with again
7		ne	îş	im	de	girt	in	e	ewe		I was not getting them also to/with again
8	heł	îş	im	ne	de	girt	in	e	ewe		I was not lifting them also to/with again

The emphatic clitic =*îş* exists as *jî* in Kurmanji, *zî* in Zazaki, =*îç* in Gorani and -*am* in Farsi.

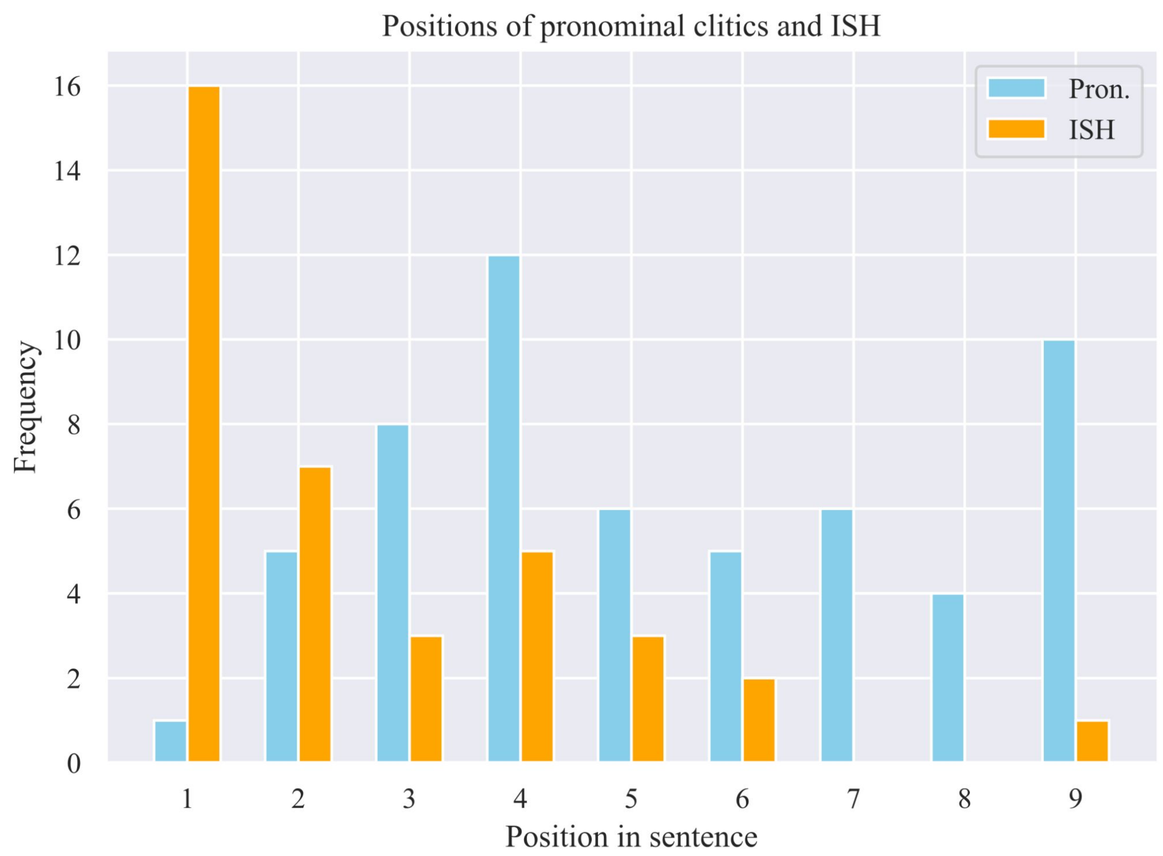
Our findings show that both endoclitics tend to attach to the leftmost morpheme within verb forms.

Analysis

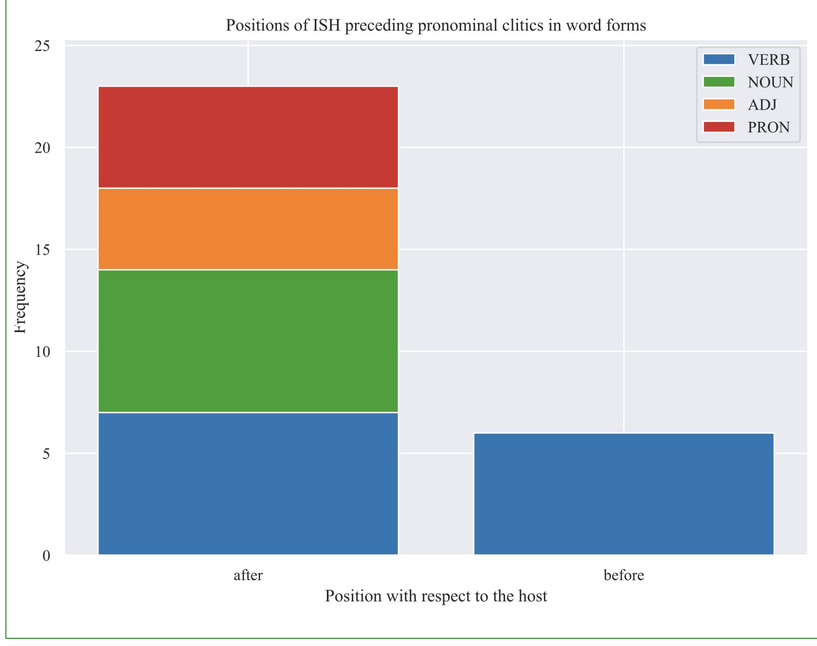
Q1: Where do the target clitics tend to appear more within a sentence?



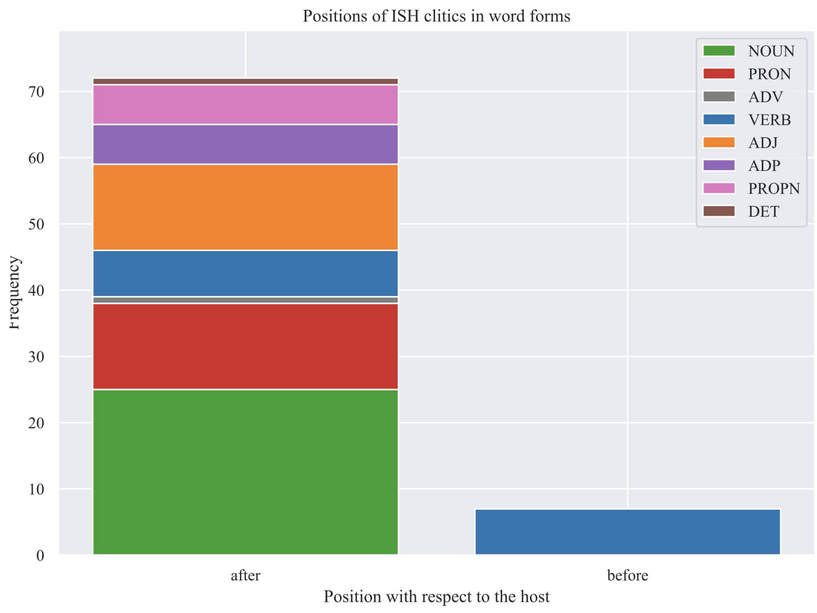
Q2: What hosts do the target clitics appear more with?

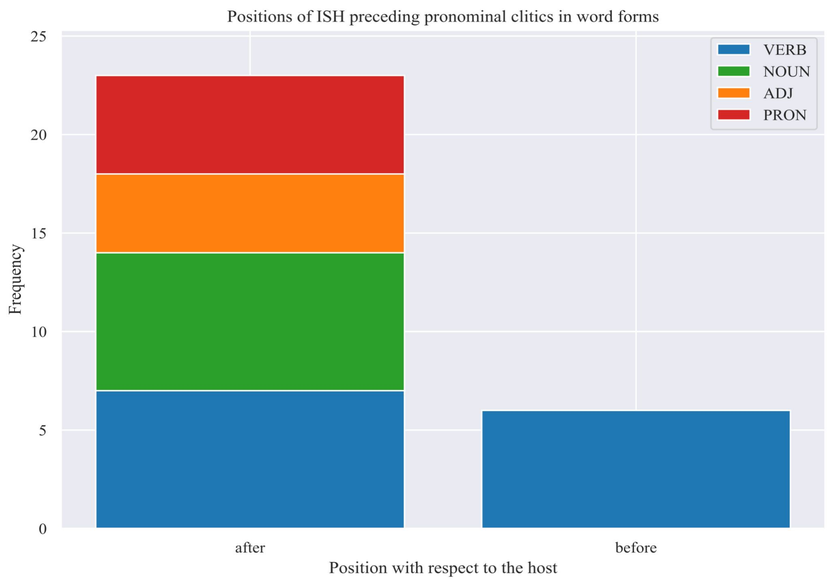


Q3: Where do the target clitics appear with respect to the host in a word?



Q4: Where do the target clitics appear with respect to the host and to each other?





Takeaways

A. While =*îş* attches to the leftmost morpheme in verbs, regardless of transitivity, pronominal endoclitics behave similarly only in transitive verbs.
B. The placement of =*îş* within a phrase varies based on the intended meaning or emphasis on a specific word.
C. The frequency of =*îş* is affected by the complex patterns of pronominal endoclitics and interactions with the personal markers.
D. The target clitics in this study appear mostly as enclitics in other varieties of Kurdish and with simpler patterns.

References

1. Bonami, O. and Samvelian, P. (2008). Sorani Kurdish person markers and the typology of agreement. In *13th International Morphology Meeting*
2. Smith, Peter W. "Non-peripheral cliticization and second position in Udi and Sorani Kurdish." Paper under revision at Natural Language and Linguistic Theory (2014).
3. Walther, G. (2012). Fitting into morphological structure: accounting for Sorani Kurdish endoclitics. In *Mediterranean Morphology Meetings,* volume 8, pages 299–321.