



NooJ dictionary for the Rromani language toward a NooJ-relevant sorting of morphosyntactic tags

Masako Watabe
Paris-Sorbonne University
ED V-433 Concepts et Langages
EA 4509 Sens, Texte, Informatique, Histoire (STIH)

NooJ Conference 2017
May 18–20, 2017
Ibn Tofail University, Kenitra, Morocco



Outline

Rromani language (I rromani ćhib)
NooJ module for the Rromani language
Rromani dialects
Categorization of Rromani dialects for NooJ
Rromani morphosyntaxe
Postpositions
Grammar .nom for postpositions
Categorization of Rromani morphosyntaxe for NooJ
Uncountable nouns in Rromani
Conclusion





Rromani language (i rromani čhib)



- Language of the Rromani people.
- Population : 10-12 million in Europe and 3 million across the Atlantic.
- Speakers : 6-8 million.
- Indo-Aryan family.
- Mostly Indian roots, but also Greek, Persian, Georgian, Armenian roots etc.
- Inflectional language; and agglutination for postpositions.

NooJ module for the Rromani language



- Polylectal; a single and common module for the four dialects of the Rromani (O-bi, O-mu, E-bi, E-mu).
- Mutual understanding is possible between good speakers of different dialects & various regions.
- Diasystem; systematic correspondence of interdialectal equivalents (i.e. diasynonym).
- How to reconstruct the Rromani diasystem as module?
- Lexical resource : polylectal & multilingual dictionary *Morri angluni rromane čhibăqi evroputni lavustik* (My first European dictionary of the Rromani language), Courthiade, 2009.



Rromani dialects

	O superdialect	E superdialect
Without phonetic mutation (bi)	O-bi	E-bi
With phonetic mutation (mu)	O-mu	E-mu



Categorization of Rromani dialects for NooJ

- Dialectal properties for NooJ :
 - ✓ 2 superdialects (O, E). \Rightarrow rro, rre
 - ✓ 2 subgroups (without, or with mutation). \Rightarrow rrbi, rrmu
- Each dialect should be represented by a combination of 2 properties.

O-bi \Rightarrow rro+rrbi

E-bi \Rightarrow rre+rrbi

O-mu \Rightarrow rro+rrmu

E-mu \Rightarrow rre+rrmu



Postpositions

- The initial Rromani module considers the postpositions to be inflectional endings.
- 4 invariable postpositions.
 - ✓ -qe : dative =>dat *for*
 - ✓ -ça : instrumental =>ins *with*
 - ✓ -θar : ablative =>abl *from*
 - ✓ -θe : locative =>loc *at*
- 1 variable postposition.
 - ✓ -qo/-qi/-qe : possessive =>poss *of*



Postpositions

o the.M.SG.DRT	ruv wolf.M.SG.DRT	<i>the wolf</i> (direct)
e the.M.SG.OBL	ruv-es wolf-M.SG.OBL	<i>the wolf</i> (oblique)
e the.M.SG.OBL	ruv-es-qe wolf-M.SG.OBL-DAT	<i>for the wolf</i>
e the.M.SG.OBL	ruv-e-ça wolf-M.SG.OBL-INS	<i>with the wolf</i>
	*ruv-es-ça	
e the.M.SG.OBL	ruv-es-θar wolf-M.SG.OBL-ABL	<i>from the wolf</i>
e the.M.SG.OBL	ruv-es-θe wolf-M.SG.OBL-LOC	<i>at the place of the wolf</i>



Postpositions

ruv-**es**-q-o

wolf-M.SG.OBL-POSS-dM.dSG.dDRT

dand

tooth.M.SG.DRT

the tooth of the wolf

ruv-**es**-q-e

wolf-M.SG.OBL-POSS-dM.dPL.dDRT

dand-**a**

tooth-M.PL.DRT

the teeth of the wolf

dand-**a**

tooth-M.PL.DRT

ruv-**es**-q-ere

wolf-M.SG.OBL-POSS-dM.dPL.dDRT

the teeth of the wolf

dand-**en-ça**

tooth-M.PL.OBL-INS

ruv-**es**-q-eren-ça

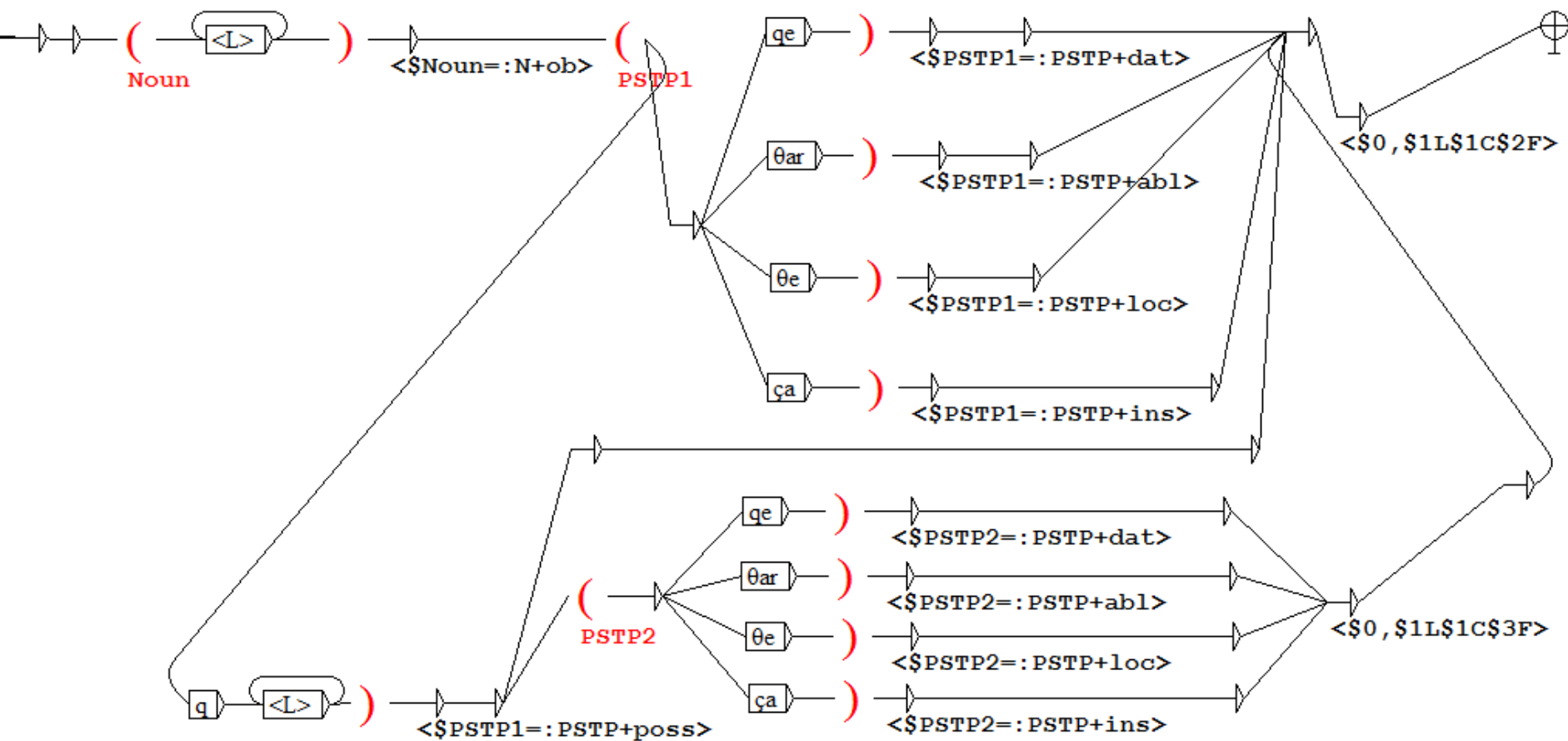
wolf-M.SG.OBL-POSS-dM.dPL.dOBL-INS

with the teeth of the wolf

Grammar .nom for postpositions

C:\Users\Masako\Documents\NooJ\rr\Lexical Analysis\2017-05_PSTP-AGG13.nom

Rromani morphological grammar.





Uncountable nouns in Rromani

- Types of nouns based on the numbers.
- ✓ Countable (singular, plural)
ruv wolf vs ruva wolfs, dand tooth vs danda teeth
- ✓ Only plural
lima snot, pixa eye-rheum
- ✓ Uncountable
love money, pani water, mas meat



Uncountable nouns in Rromani

Okote si jekh ruv.

A wolf is there.

Okote si but ruva.

Many wolfs are there.

Si man but love.

I have a lot of money

Piav but pani.

I drink a lot of water.

Xav but mas.

I eat a lot of meat.

Akate si jekh lovorro.

One coin is here.

Xalom marikli maseça.

I ate pie with meat.

Xalom marikli masença.

I ate pie with mixed meat.

Buuuut masa si akate!

Here is a looooot of meat!



Conclusion

- To import the published dictionary to NooJ dictionary, we have to advance diasystem study and then to check the coherence at dialectological & technical levels.
- The grammar .nom for postpositions is a test pattern and has not been completed.
- About uncountable nouns, we could define some more types of specificities.
- And show must go on!!



NooJ dictionary for the Rromani language toward a NooJ-relevant sorting of morphosyntactic tags



References

Courthiade Marcel et alii, *Morri angluni rromane čhibăqi evroputni lavustik* (My first European dictionary of the Rromani language), Cigány Ház Budapest, 2009.

R.E.D.-RRROM (Restoring the European Dimension of Rromani Language and Culture), 2014-, available on www.red-rrom.com

Silberztein Max, *La formalisation des langues : l'approche de NooJ*, ISTE Eds., Londres, 2015.



**NooJ dictionary for the Rromani language
toward a NooJ-relevant sorting of morphosyntactic tags**

Te oven saste veste!

Thank you!

Masako Watabe, masakowatabe@free.fr

