

Extracting verb valency frames with NooJ

Krešimir Šojat, Kristina Vučković*, Marko Tadić

ksojat@ffzg.hr, kvuckovi@ffzg.hr, marko.tadic@ffzg.hr

Faculty of Humanities and Social Sciences

University of Zagreb

Department of Linguistics

*Department of Information Sciences

Ivana Lucica 3, Zagreb, Croatia



The Plan

- **Our agenda?**
 - Full description of consumption verb valency frames (FrameNet by Fillmore, Atkins, Ruppenhofer et al, etc.)
 - given core arguments
 - searching for peripheral elements
 - time, place, manner, company (PP+I), instrument (NP+I), cause...
- **How?**
 - using core verb valency frames description
 - checking the verb's environment
 - -4 and +4 sets of word phrases
- **Why?**
 - to prepare data for Croatian WordNet
 - to improve grammars for syntactic verb environment recognition



Overview

- Croatian consumption verb valency
 - main characteristics
- Lexicon
 - data description
- Syntactic grammar
 - detecting verb's environment
- Checking the data
 - extracting frames



Consumation verb valency lexicon

- adding semantic information to lexicon

- semantic field = cons

- consumer

- cons1 (Nominative)

Ja jedem. (I am eating.)
Jedem. (I am eating.)

- consumed

- cons2 (Genitive)

Ona se najela gljiva. (She has stuffed herself with mushrooms.)

- cons4 (Accusative)

Ja jedem ribu. (I'm eating fish.)

- cons7 (Instrumental)

Oni se hrane kukuruzom.

- core arguments = cons1

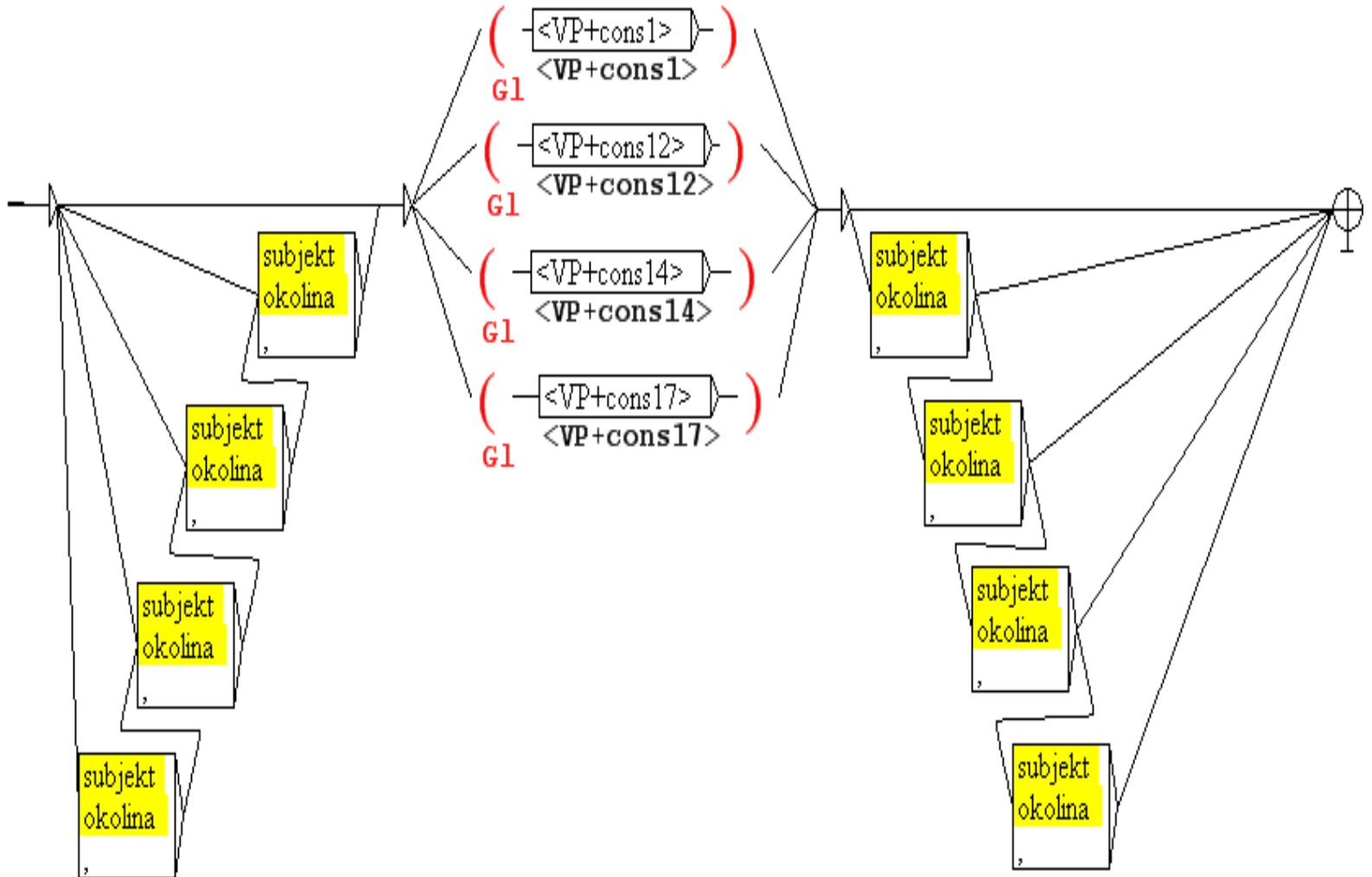
(They are feeding on corn.)

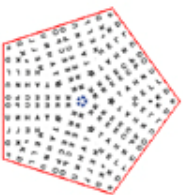
| cons12 | cons14 | cons17

- jesti, V+FLX=JESTI+Aspect=inf+Prelaz=pov

+cons+cons1+cons14

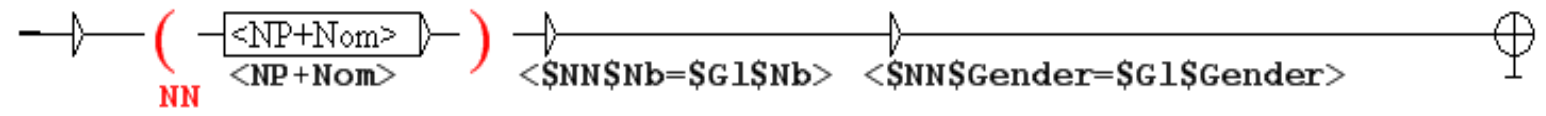
Grammars



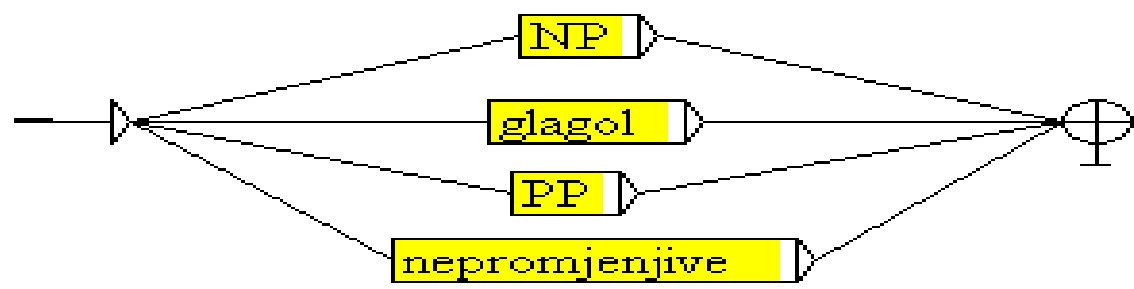


Grammars

subjekt



okolina





Results

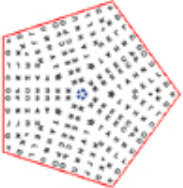
Kao

**i većina drugih, ta obitelj nikad ne jede u Branimirovoj
već hranu nosi
kući.**

Like

**many others, that family never eats in Branimirova street
but carries their food
home.**

-4	-3	-2	-1	0	1	2	3	4
i	većina drugih	ta obitelj	nikad	ne jede	u Branimirovoj	već	hranu	nosi
<C>	<NP+N om>	<NP+Nom>	<R>	<VP+cons1>	<PP+L>	<C>	<NP+Acc>	<VP>



Results 2 : problems

- **A:** Ona se tako hrani poradi svoga siromaštva što ga ne smije otkriti kćeri.

-4	-3	-2	-1	0	1	2	3	4
	ona	se	tako	hrani	poradi svoga siromaštva	što	ga	ne smije otkriti
	<NP+Nom>	<VP>	<R>	<VP+cons1>	<PP+G>	<PRO>	<NP+Acc>	<VP>

- 
- **B:** Prije početka susreta jeli su kroasane i voće i pili voćne sokove.

-4	-3	-2	-1	0	1	2	3	4
			prije početka susreta	jeli su	kroasane i voće	i	pili	voćne sokove
			<PP+G>	<VP+cons14>	<NP+Acc>	<C>	<VP+cons14>	<VP>



Possible solutions 1

- A:
 - <VP+cons1><PP+G><PRO+question><WF>
- =>
 - <VP+cons1> <ADV+cause> <PP+G> <Att> > >

- A: <PP+G> - ADV+cause
- B: <PP+G> - ADV+time (S+vr)
 - <PP+G><VP+cons14>...
- =>
 - <ADV+time> <PP+G> > <VP+cons14>

Possible solutions 2

- A:** Ona se tako hrani poradi svoga siromaštva što ga ne smije otkriti kćeri.

ona	-2	tako	0	poradi svoga siromaštva što ga ne smije otkriti kćeri.
	se	<ADV v +ma nne r>	hrani	
<NP+ CONSUMER>	<VP>		<VP+cons1>	<ADV+cause>

- B:** Prije početka susreta jeli su kroasane i voće i pili voćne sokove.

-4	-3	-2	prije početka susreta	0	kroasane i voće	3	4
				jeli su		pili	voćne sokove
			<ADV+time>	<VP+cons14>	<NP+CONSUME D>	<VP+cons14>	<VP> 11/12



Future work

- building local grammars for recognizing
 1. syntactic verb valency frames
 - morphosyntactic description of phrases
 2. semantic verb valency frames
 - core + peripheral frame elements
 3. check if described frames can be copied into other semantic fields