

SynCro - Parsing Simple Croatian Sentences

Kristina Vučković, Božo Bekavac, Zdravko Dovedan
University of Zagreb, Faculty of Humanities and Social Science
{kvuckovi, bbekavac, zdovedan}@ffzg.hr

NooJ2009
Tozeur
2009-06-09



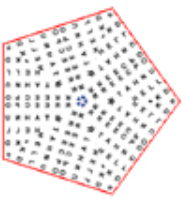
Overview of the work

- on our way to complete parse trees for Croatian sentences
 - building partial trees for simple Croatian sentences
 - favor of precision over recall
- Building syntactic parse trees for Croatian is quite a challenge!
- properties of Croatian language:
 - mostly free word (better to say constituent?) order
→ scrambling very frequent
 - could be a long distance between parts of predicate

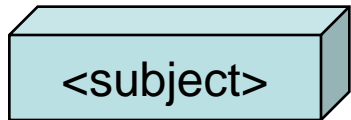
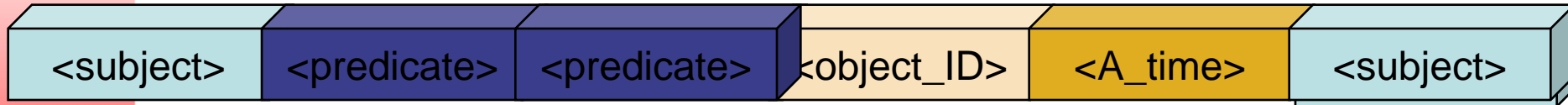


Syntactic trees

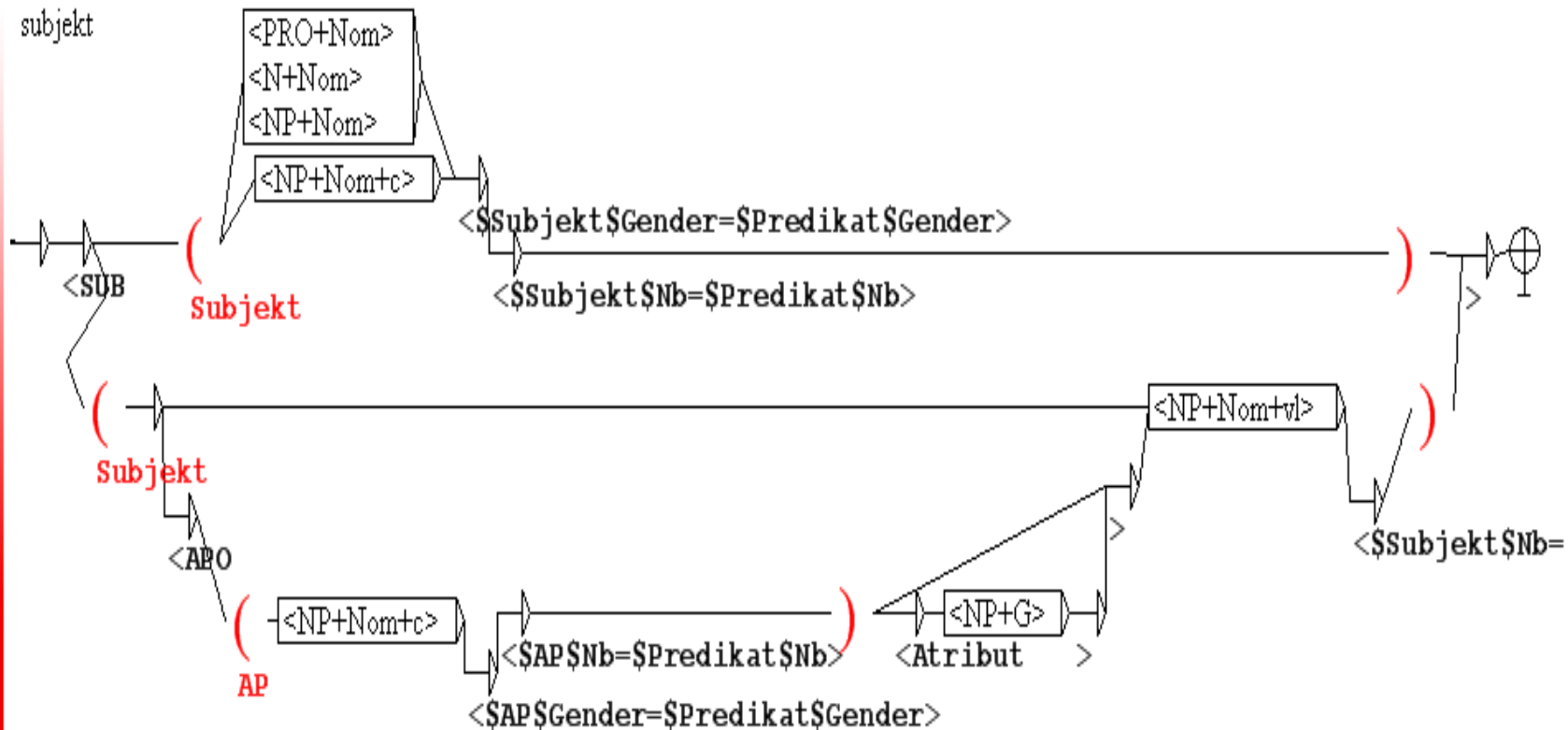
- at this stage, the following structures of parse trees are described
 - verb
 - subject - verb
 - subject - verb - direct object
 - subject - verb - direct (and, or indirect object)
- in addition to all these combinations, we can detect
 - adverbial phrases of place and
 - adverbial phrases of time (temporal expressions)
- Complex combinations (e.g. multiple PPs) are not described



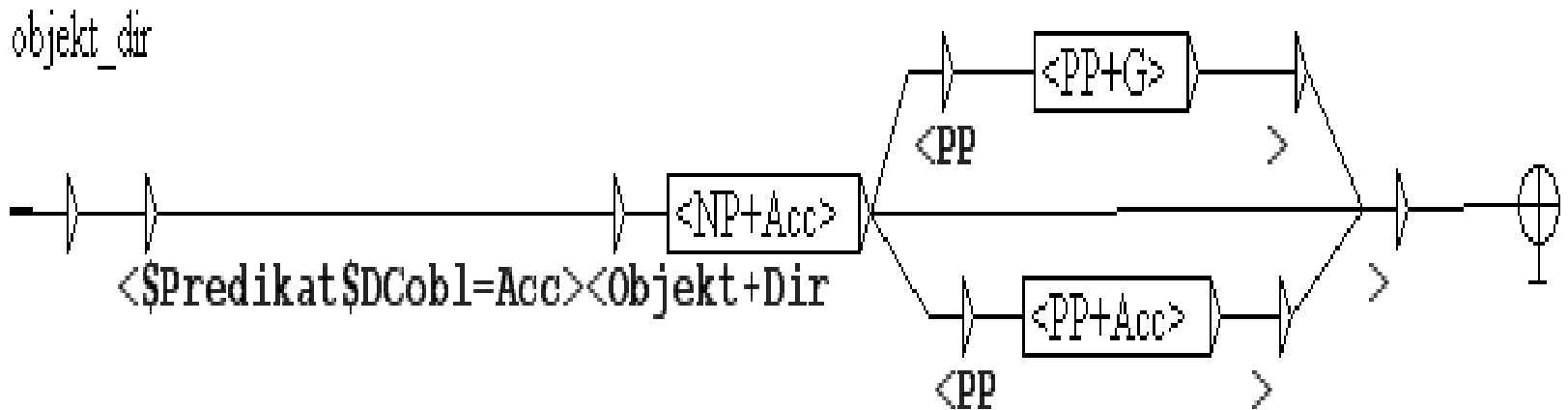
Sentence structure in Croatian



Recognition of: subject node

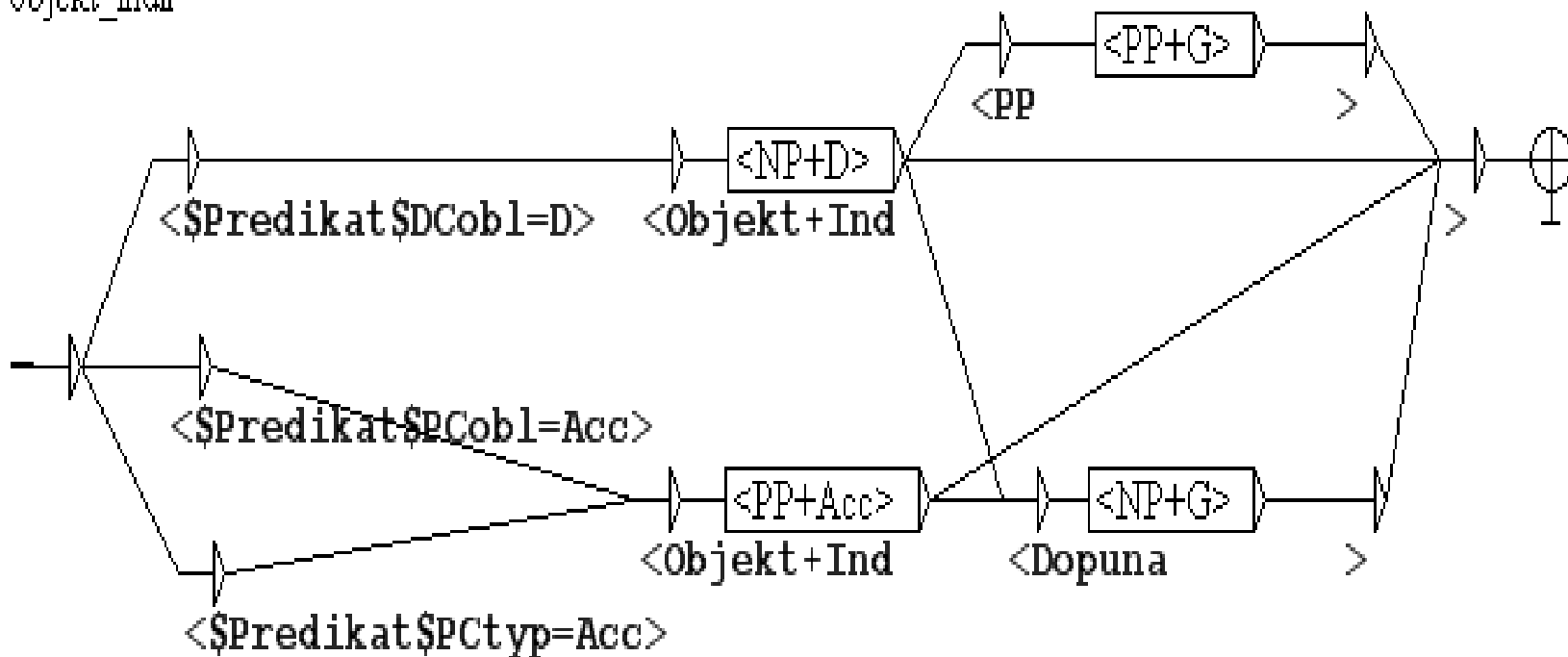


Recognition of: direct object node



Recognition of: indirect object node

objekt_indir

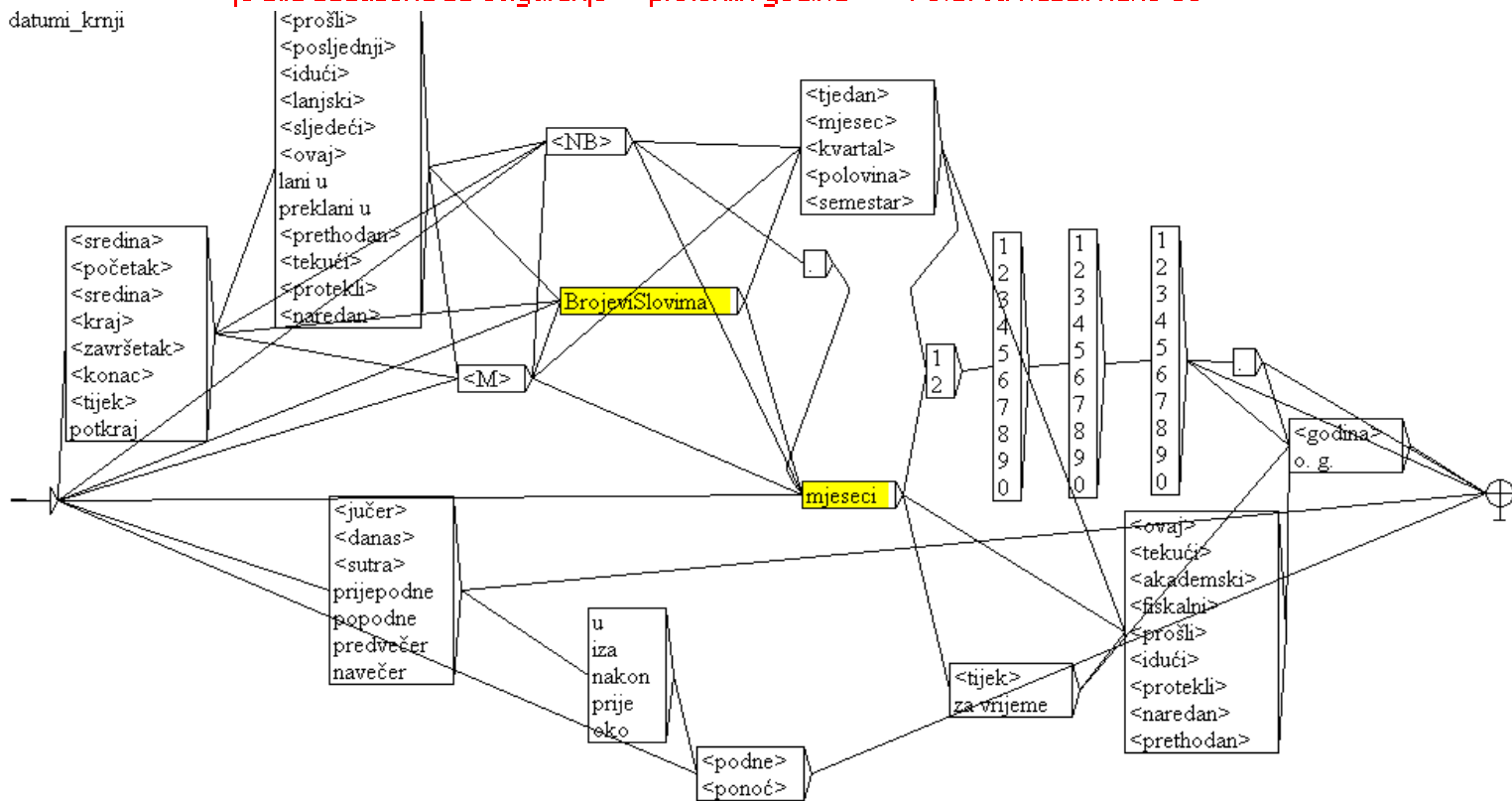


Recognition of: Adverbs of Time

Bosni i Hercegovini (BiH), u
i Hercegovini (BiH), u srijedu (11.
žrtava pokolja koje su identificirane
komemoraciju postale su sporna točka
je bila zadužena za osiguranje

srijedu
11. srpanj
prošle godine
prošlog mjeseca
proteklih godina

(11. srpanj).
)
. .
<p>Godišnjicu srebreničkog pokolja
<p>Godišnjicu srebreničkog pokolja
<p>Sigurnosne mjere za
kada je Schwarz-Schilling propočio
. Srbi su kazali kako će

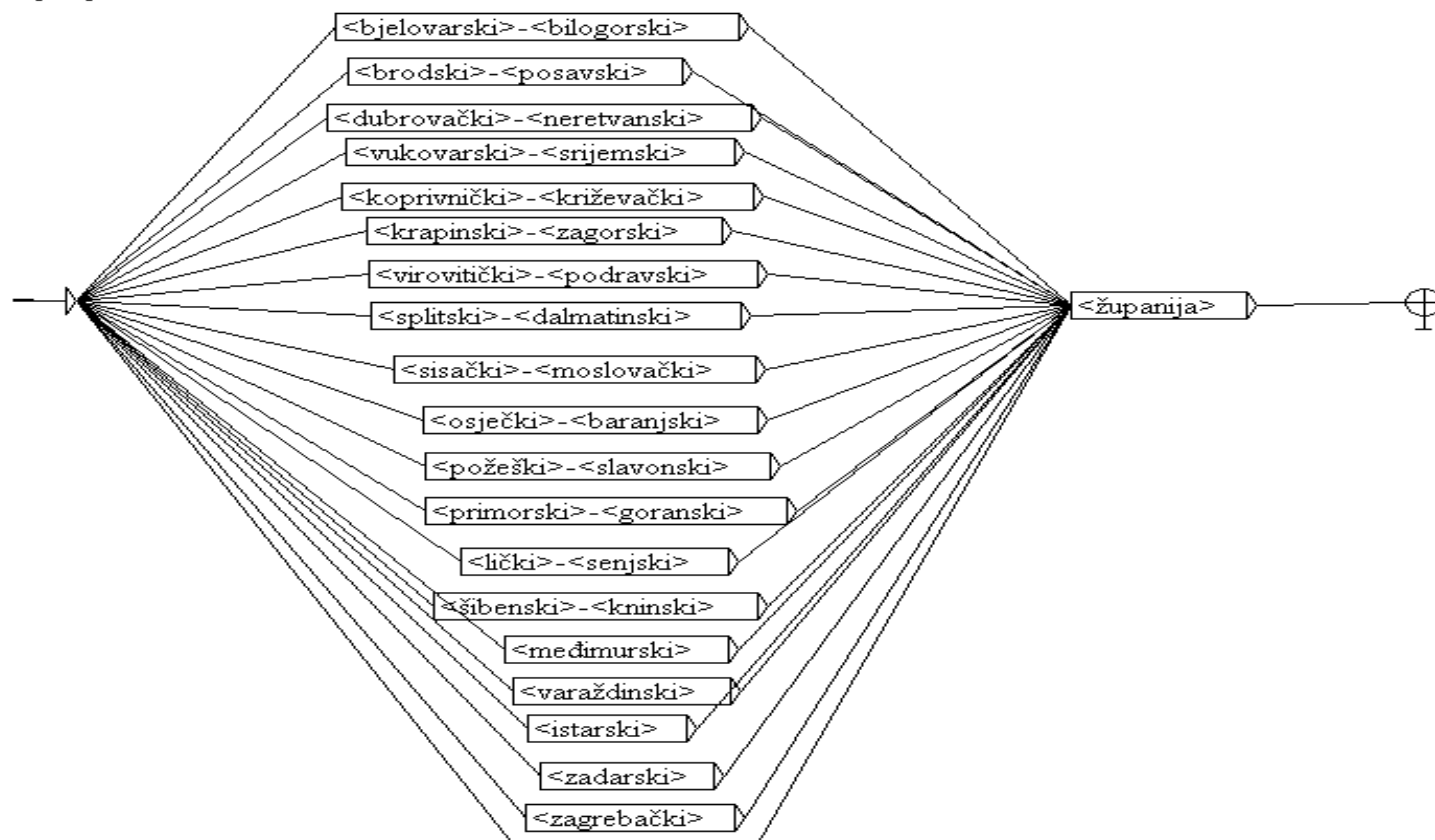


Recognition of: Adverbs of Place

NATO-u prolazi kroz sporazum s Grčkom
zemalja, sa sjedištem u Lyonu u Francuskoj
-a optuženi za krijumčarenje ljudi u Makedoniju
koji pokazuju da su ušli u Makedoniju
kubičnih metara ruskog zemnog plina u Italiju

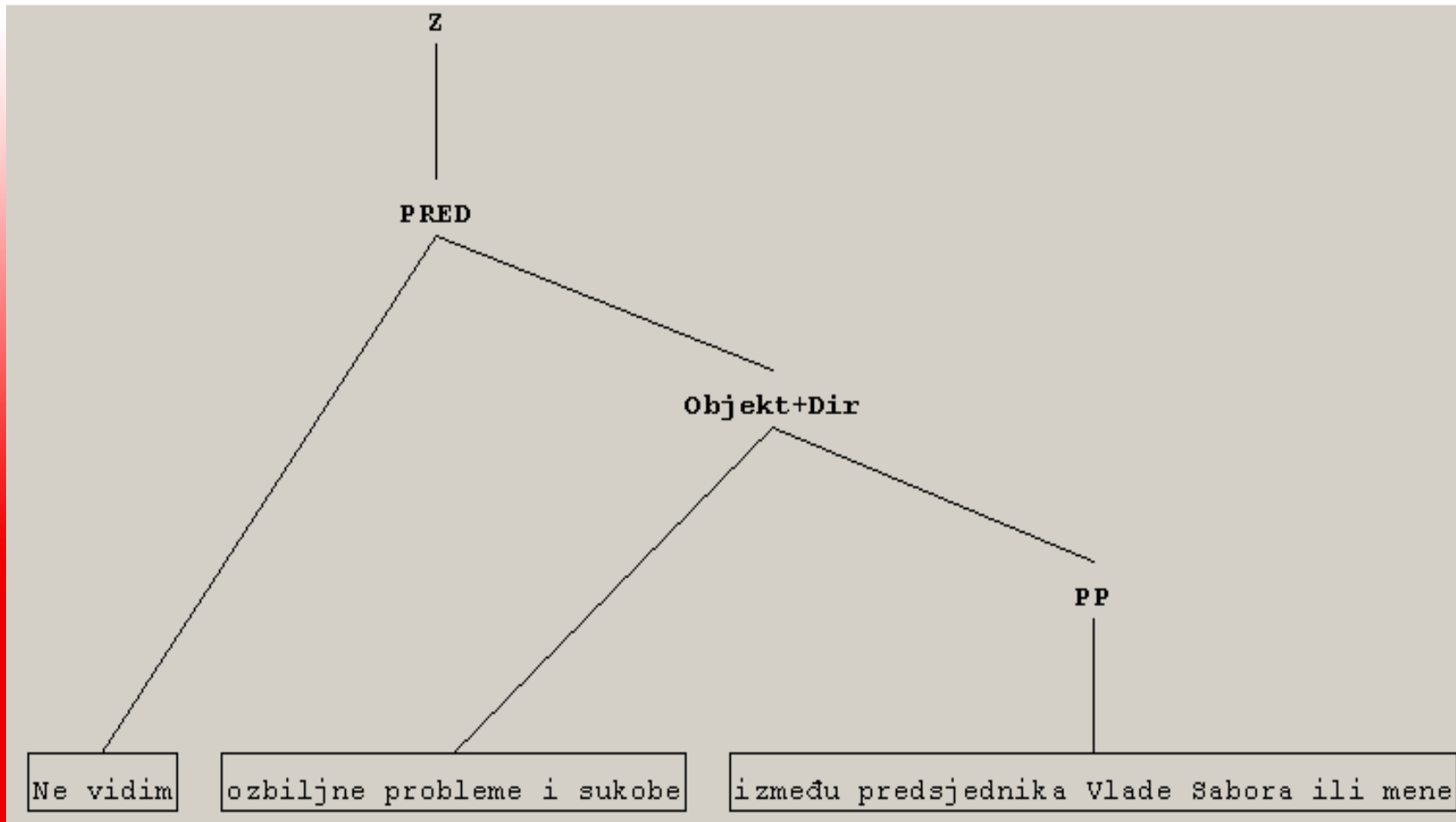
oko njezina naziva".
, također je postavila četiri rekonstruirane
Dužnosnici kažu kako su oni
, odnosi kao prema pitanju od

zupanije



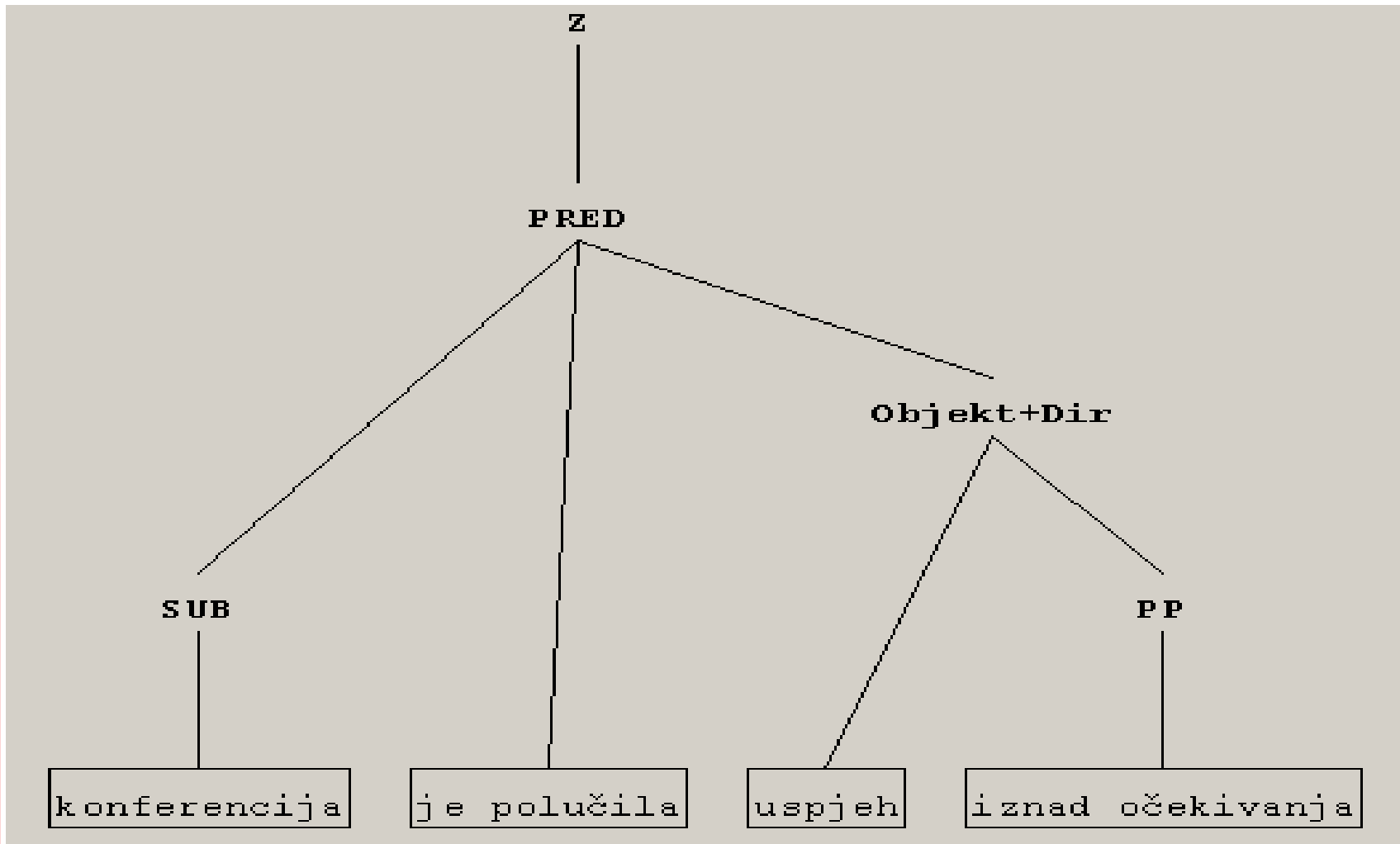
Examples

Ne vidim ozbiljne probleme i sukobe između predsjednika, Vlade, Sabora ili mene.



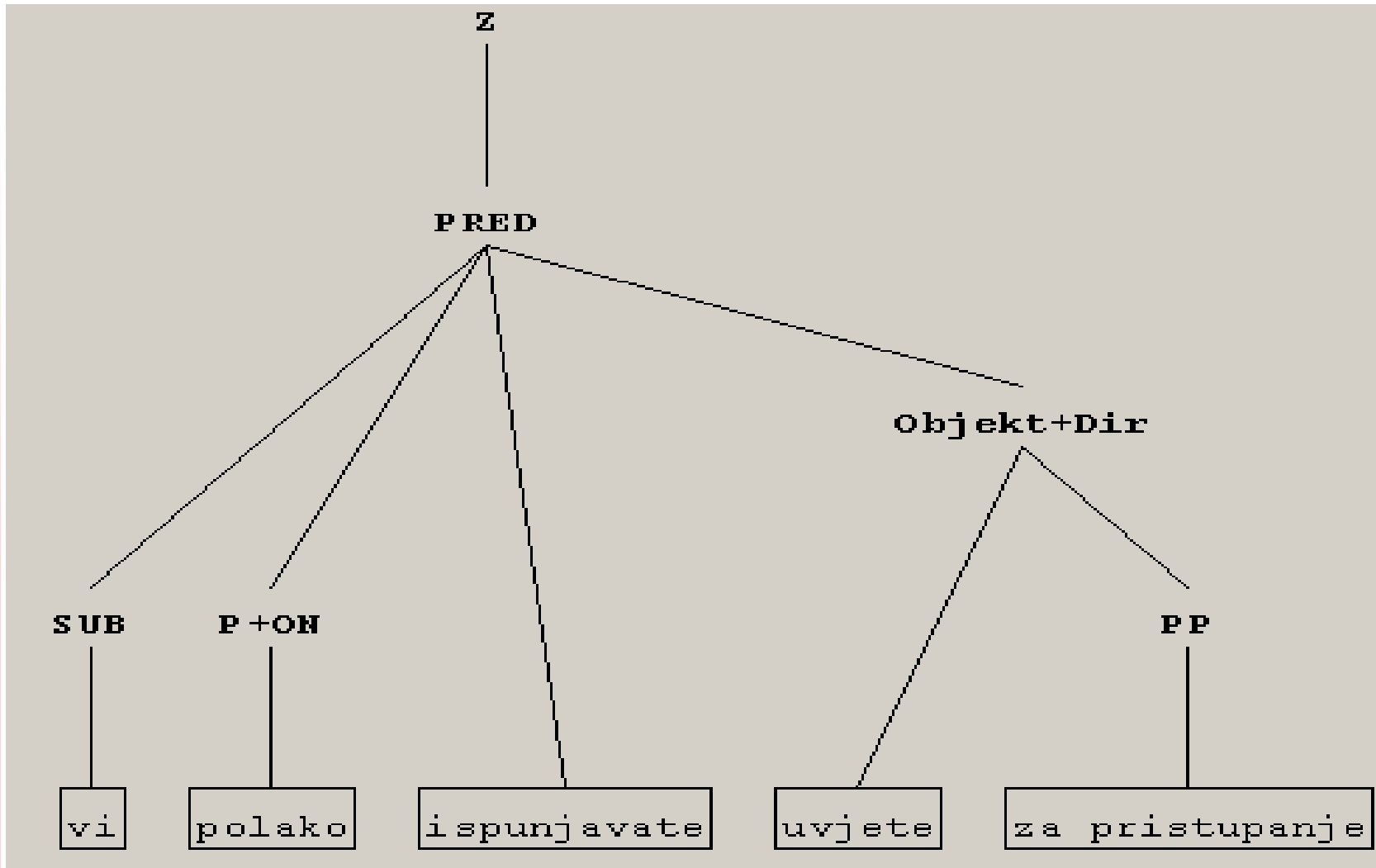
Examples

Konferencija je polučila uspjeh iznad očekivanja.



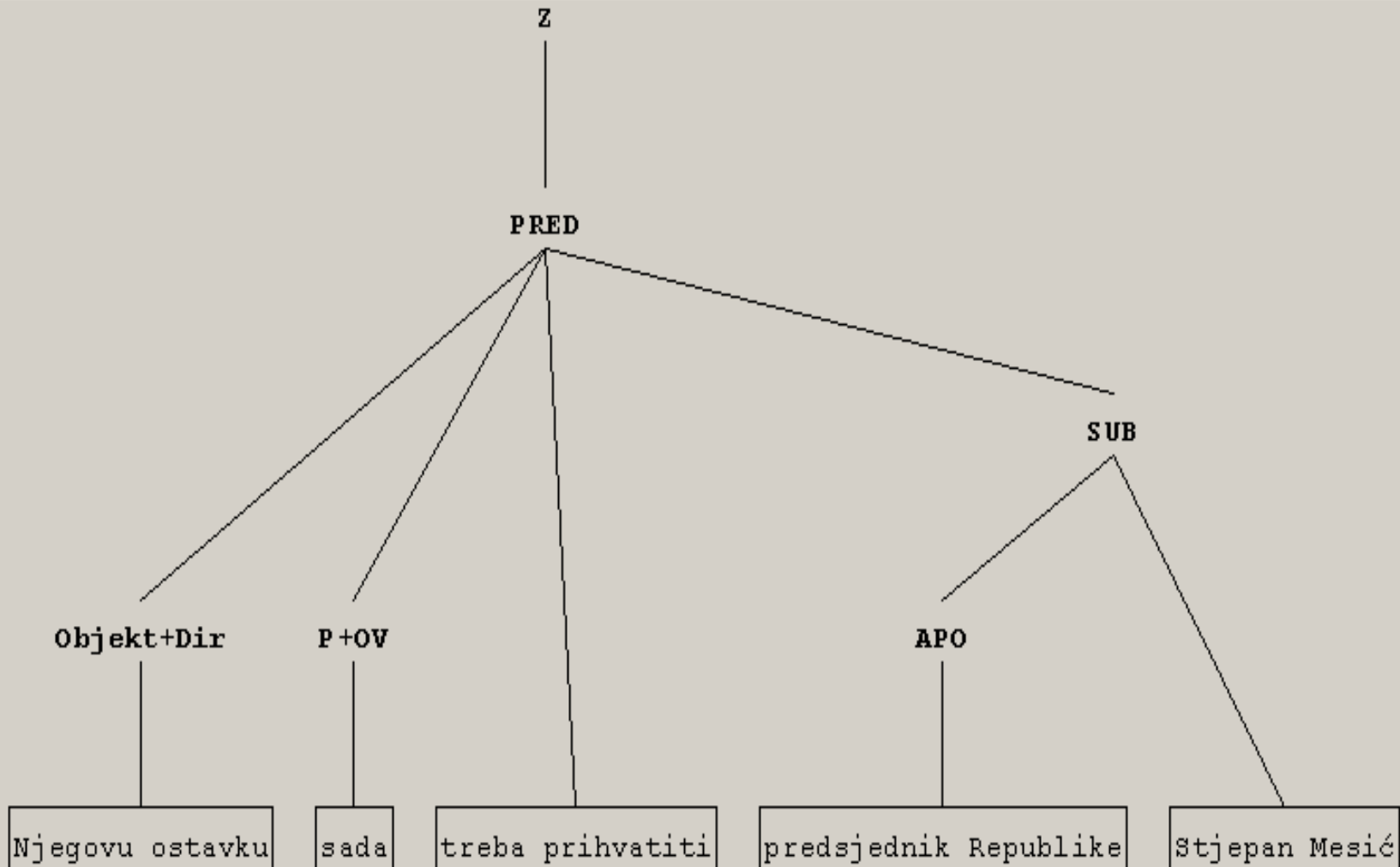
Examples

Vi polako ispunjavate uvjete za pristupanje.



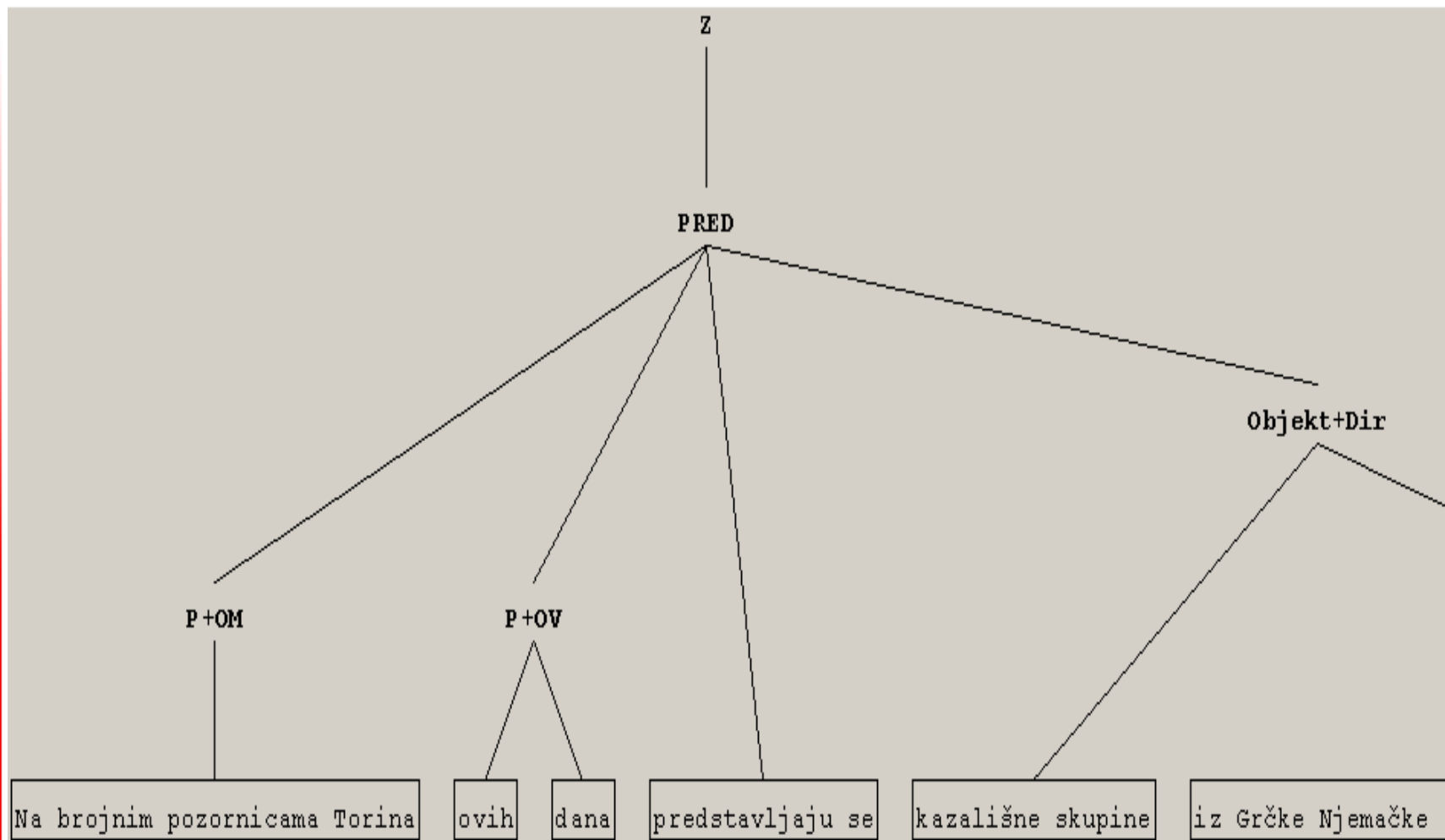
Examples

Njegovu ostavku sada treba prihvatiti predsjednik Republike Stjepan Mesić.



Examples

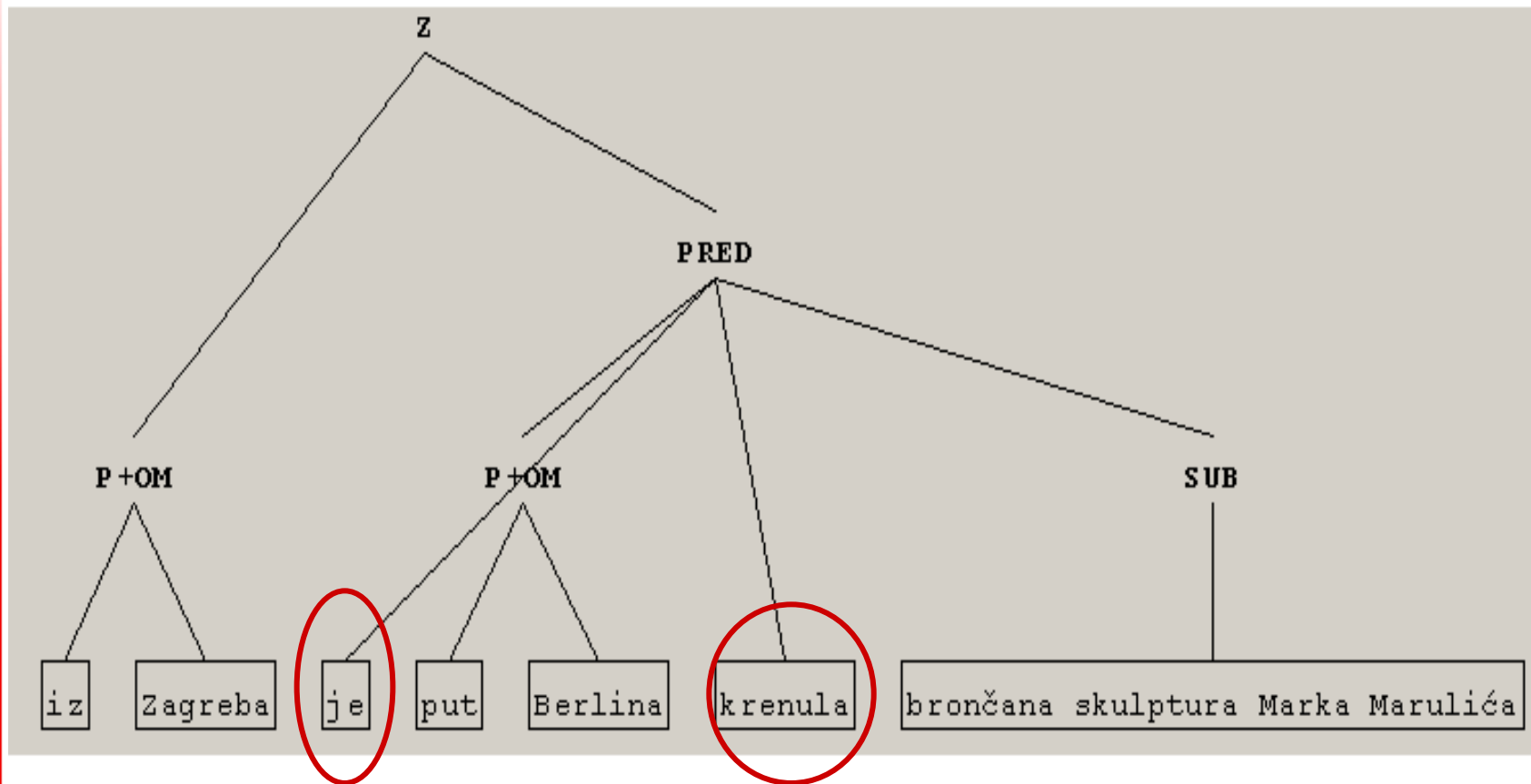
Na brojnim pozornicama Torina ovih dana predstavljaju se kazališne skupine iz Grčke, Njemačke, Italije, Latvije



Examples

- dislocated predicate <PRED>

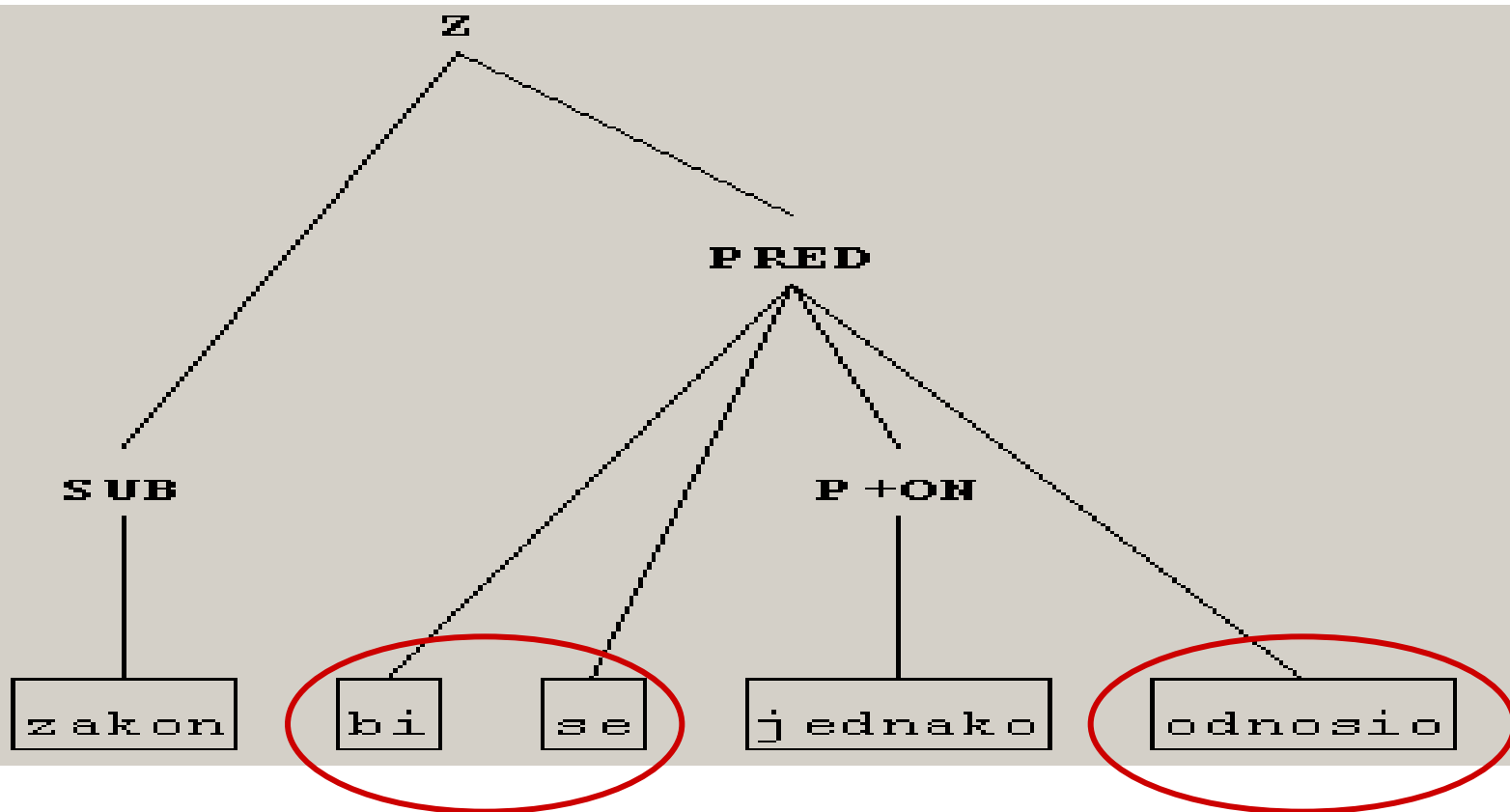
Iz Zagreba je put Berlina krenula brončana skulptura Marka Marulića.



Examples

- dislocated predicate <PRED>

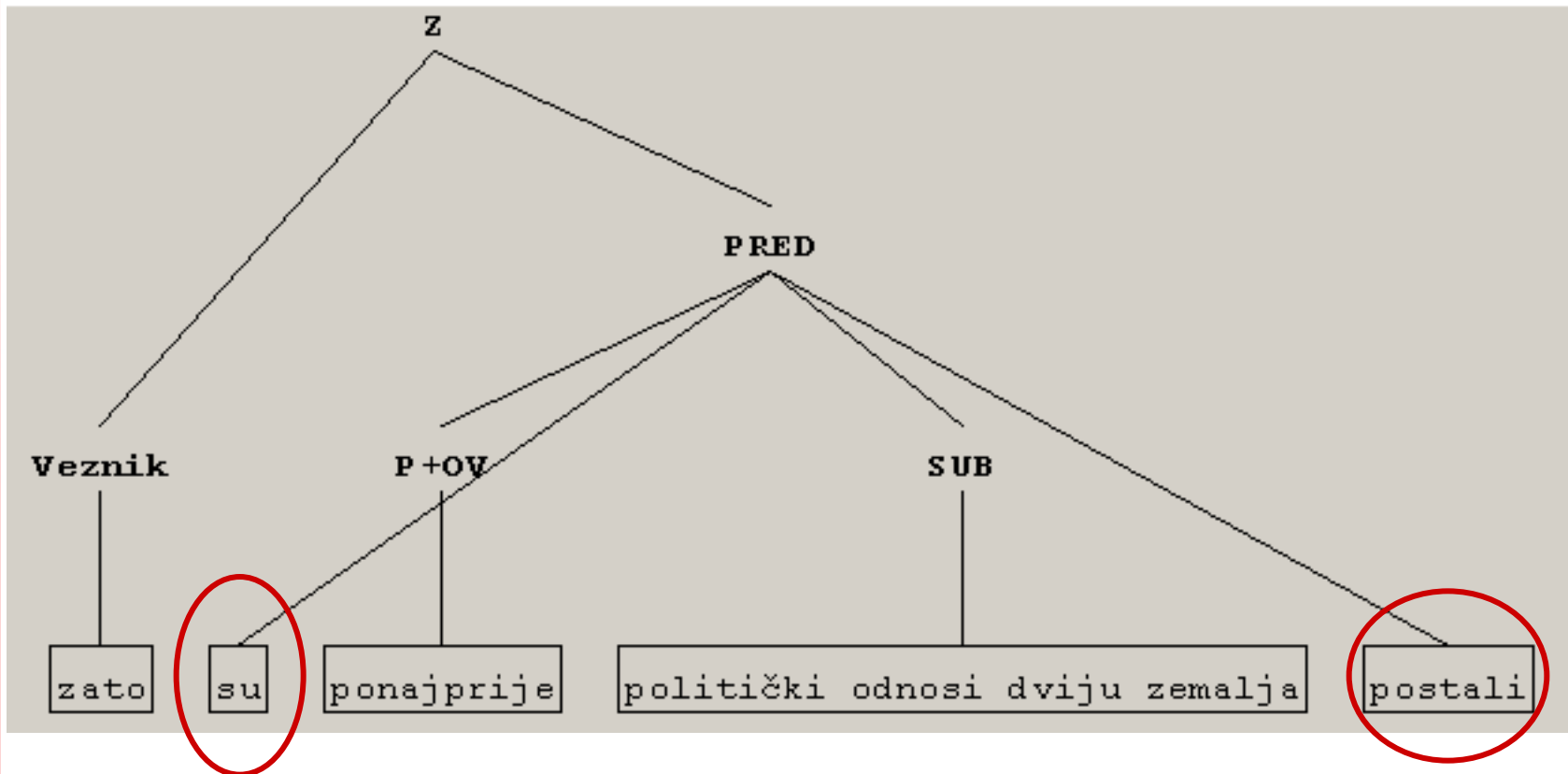
Zakon bi se jednako odnosio



Examples

- dislocated predicate <PRED>

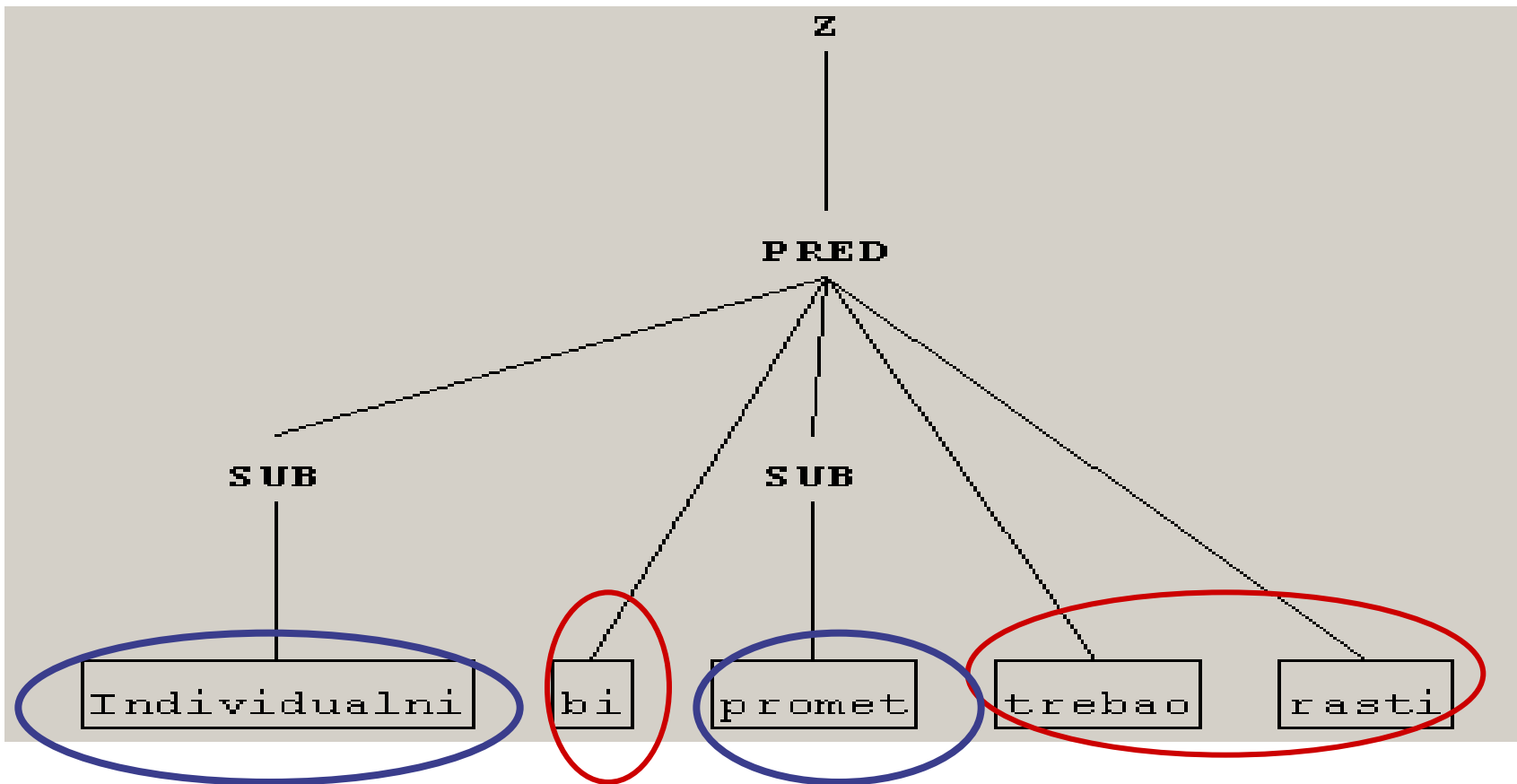
Zato su ponajprije politički odnosi dviju zemalja postali

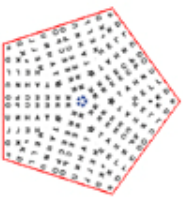


Examples

- dislocated predicate <PRED> and dislocated subject <SUB>

Individualni bi promet trebao rasti





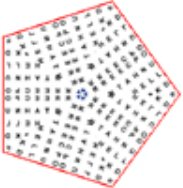
Addition to the Lexicon

- For recognizing Adverbials of Time and Place:
 - Adjectives +time / + place
 - R+vr
 - R+mt
 - Prepositions +time / + place
 - S+vr
 - S+mt



Problems and Future Work

- if there is a nominal predicate
 - ex.
 - *The child was very smart.*
- if the <NP> node is a coordination made of 2 or more <NP>'s of different gender
 - > *The child was*
 - ~~ex.~~ *was very smart.*
 - *The boy and a girl were very smart.*
 - > *The boy and a girl*
 - ~~ex.~~ => <NP+p+m>
 - *He was singing from March to April.*
 - *He was singing from Paris to Tozeur.*
- if <S> has more than one annotation



Syntactic Analysis

- 1 + Concordance Entry / 1

Display: Structural Tree Derivation Tree

Lexeme, Lemma, Linguistic Information

Thank You for Your Attention!

```
graph TD; S[S] --- PRE[PRE]; S --- Obj_D1[Obj_D]; PRE --- Obj_ID[Obj_ID]; PRE --- Obj_D2[Obj_D]; Obj_ID --- Thank[Thank]; Obj_D2 --- for[for]; Obj_D2 --- Your[Your]; Obj_D2 --- Attention[Attention]; for --- exclamation[!];
```