

Treebank querying with GrETEL 3

bigger, faster, stronger

Liesbeth Augustinus, Bram Vanroy & Vincent Vandeghinste
Centre for Computational Linguistics, KU Leuven

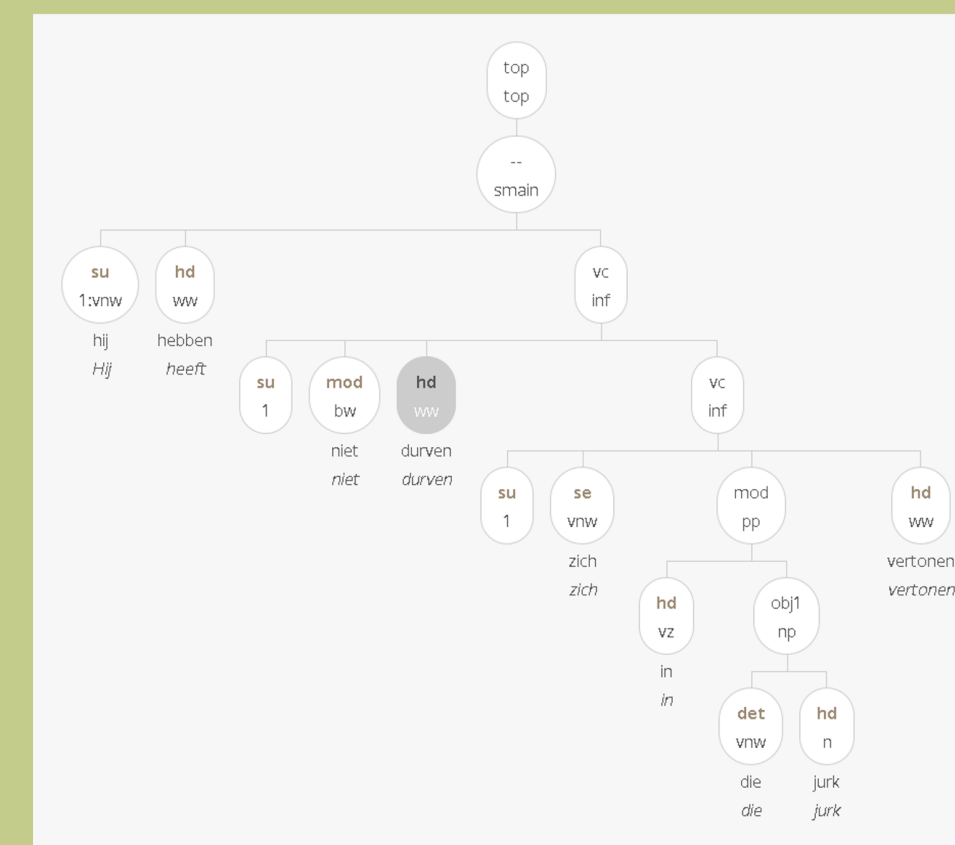


1 example

Find instances of *te* omission

"Hij heeft zich niet in die jurk durven vertonen."
(He has not dared to show up in that dress.)

new tree
visualisation



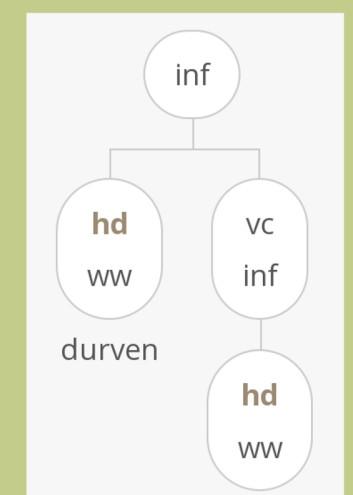
parse tree

2

3 selection matrix

Show advanced options									
sentence	Hij	heeft	zich	niet	in	die	jurk	durven	vertonen
word	o	o	o	o	o	o	o	o	o
lemma	o	o	o	o	o	o	o	*	o
word class	o	o	o	o	o	o	o	o	*
optional in search	*	*	*	*	*	*	*	o	o

dynamic subtree
generator



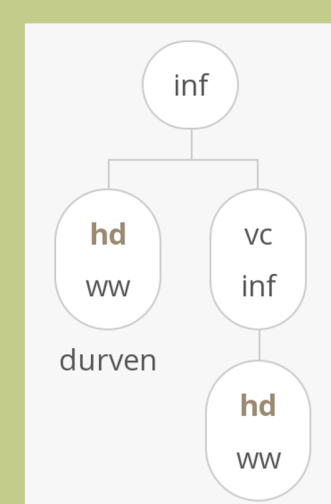
stronger

CGN	LASSY	SoNaR
Spoken Dutch	Written Dutch	Written Dutch
1M tokens	1M tokens	500M tokens
Corrected	Corrected	Not corrected

treebank selection

4

5 query



automatic
XPath generator

```
//node[@cat="inf" and  
node[@rel="hd" and @pt="ww" and  
@lemma="durven"] and  
node[@rel="vc" and @cat="inf" and  
node[@rel="hd" and @pt="ww"]]]
```

GrInding (GrETEL Indexing)
speeds up search
in SoNaR

faster

Hij scoort in de peilingen beter dan zijn medewerkers ooit hadden
durven hopen.

Maar vooral de kracht die bij deze mensen aanwezig is, dient
ondersteund te worden, waardoor ze als vanzelf een appèl zouden
moeten durven doen op de bestaande voorzieningen.

Ik zou me niet durven baseren op de mening van toevallig aanwezige
klanten in een winkelpunt.

results

6

Try it yourself at : <http://gretel.ccl.kuleuven.be>