

# **Enriching a grammatical database with intelligent links to linguistic resources**

**Ton van der Wouden, Gosse Bouma,  
Matje van de Kamp, Marjo van Koppen,  
Frank Landsbergen, and Jan Odijk**

# The Taalportaal

The Taalportaal project ([www.taalportaal.org](http://www.taalportaal.org)) is a collaboration of the Meertens Institute, the Fryske Akademy, the Institute of Dutch Lexicology and Leiden University, funded, to a large extent, by the Netherlands Organisation for Scientific Research (NWO). The project is aimed at the development of a comprehensive and authoritative scientific grammar for Dutch and Frisian in the form of a virtual language institute.

**Dutch** Frisian Afrikaans

## Phonology

Phonotactics  
Phonological processes

## Morphology

Inflection and derivation  
Compounding

## Syntax

Adjectives and Adjective Phrases  
Adpositions and adpositional phrases  
Nouns and Noun Phrases

## Welcome

*Taalportaal* will be a comprehensive and authoritative scientific grammar of Dutch, Frisian and Afrikaans, written and compiled by linguists, for linguists.



## Questions about Taalportaal:

- What is *Taalportaal*?
- What is the target audience of *Taalportaal*?
- What variants of Dutch are covered in *Taalportaal*?

► [More questions ...](#)

## Highlighted topic:

*As NP argument of the verb*

## By linguists, for linguists

Meet the linguistic community behind *Taalportaal*



► [Read more ...](#)

## About the project

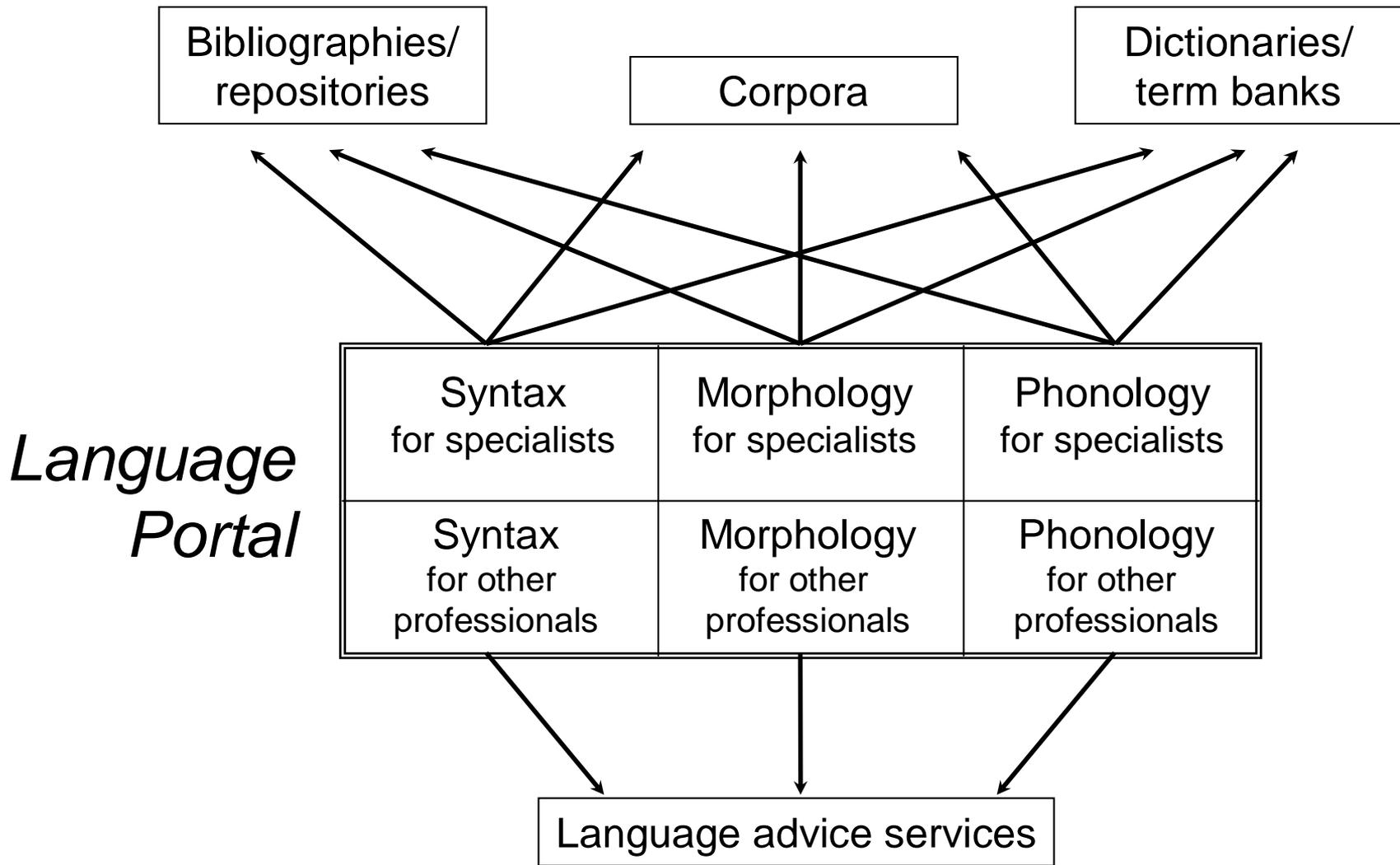
Learn about the *Taalportaal* project



► [Read more ...](#)

## Bibliography

## Glossary



## Phonology

Phonotactics  
Phonological processes

## Morphology

Inflection and derivation  
Compounding

## Syntax

Adjectives and Adjective Phrases

- 1 Characteristics and classification
- 2 Projection of adjective phrases I: Complementation
  - Introduction
  - 2.1. Prepositional complements**
  - 2.2. Nominal complements
  - 2.3. Discontinuous adjective phrases
  - 2.4. Pronominalization of the adjective (phrase)
  - 2.5. Bibliographical notes
- 3 Projection of adjective phrases II: Modification
- 4 Projection of adjective phrases III: Comparison
- 5 Attributive use of the adjective phrase
- 6 Predicative use of the adjective phrase
- 7 The partitive genitive construction
- 8 Adverbial use of the adjective phrase
- 9 Participles and infinitives: their adjectival use
- 10 Special constructions

Adpositions and adpositional phrases

Nouns and Noun Phrases

## 2.1. Prepositional complements

The examples in (2) show that complements of adjectives are normally PPs, which are often optional.

Example 2

- a. Jan is niet boos (over die opmerking).  
Jan is not angry about that remark
- b. Jan is niet tevreden (over zijn beloning).  
Jan is not satisfied with his reward

The examples in (3) further show that complements of adjectives can sometimes be clauses, which are introduced by the (often optional) **ANTICIPATORY PRONOMINAL PP** *er + P'P + it'*; given that, due to the phonological weakness of the pronominal element *er*, this PP is normally split by means of **R-EXTRACTION**, we will use italics to make it easier for the reader to detect the constituting parts of this PP.

Example 3

- a. Jan is (*er*) boos (*over*) dat Peter niet gekomen is.  
Jan is there angry about that Peter not come is  
'Jan is angry (about it) that Peter didn't come.'
- b. Jan is (*er*) tevreden (*over*) dat hij ontvangen is.  
Jan is there satisfied with that he received is  
'Jan is satisfied (about it) that he has been received.'

The option of having an anticipatory pronominal PPs in (3) indicates that these examples are related to the examples in (2). The following two subsections will discuss examples of this sort more extensively.

- [+] **I. Prepositional complements: Adjective + P-NP**
- [+] **II. Clauses introduced by the anticipatory PP-complement *er + P***
- [+] **III. A note on adjectives modified by *te*, *aenoea*, and *voldoende***

## ABOUT THIS TOPIC

**Author(s):**  
Hans Broekhuis

**Category:**  
Syntax

**Publication date:**  
01 June 2015

**Keywords:**  
Adjectival, Adjective, Anticipatory, Anticipatory PP-complement, Clause, Committative, Complementation, Control, Deverbal Adjective, Infinitival Complementation, Obligatory Complementation, Optional Complementation, PP, Participle, Past Participle, Passive Participle, Prepositional Complement, Pseudo-Participle

cite

print

This topic is the result of an automatic conversion from Word and may therefore contain errors. A free Open Access publication of the corresponding volumes of the Syntax of Dutch is available at OAPEN.org.

In his largely positive review of the SoD volumes on NP syntax in *LINGUA* (2013), Jack Hoeksema points out that “There is a growing body of work in empirical studies of judgment variation [...] that future extensions of this grammar could benefit from, especially when coupled to studies of actual usage patterns in corpus material”.

By enriching the on-line version of SoD with queries over syntactically annotated corpora, the current project tries to accommodate the needs of researchers like Hoeksema.

# Some relevant resources

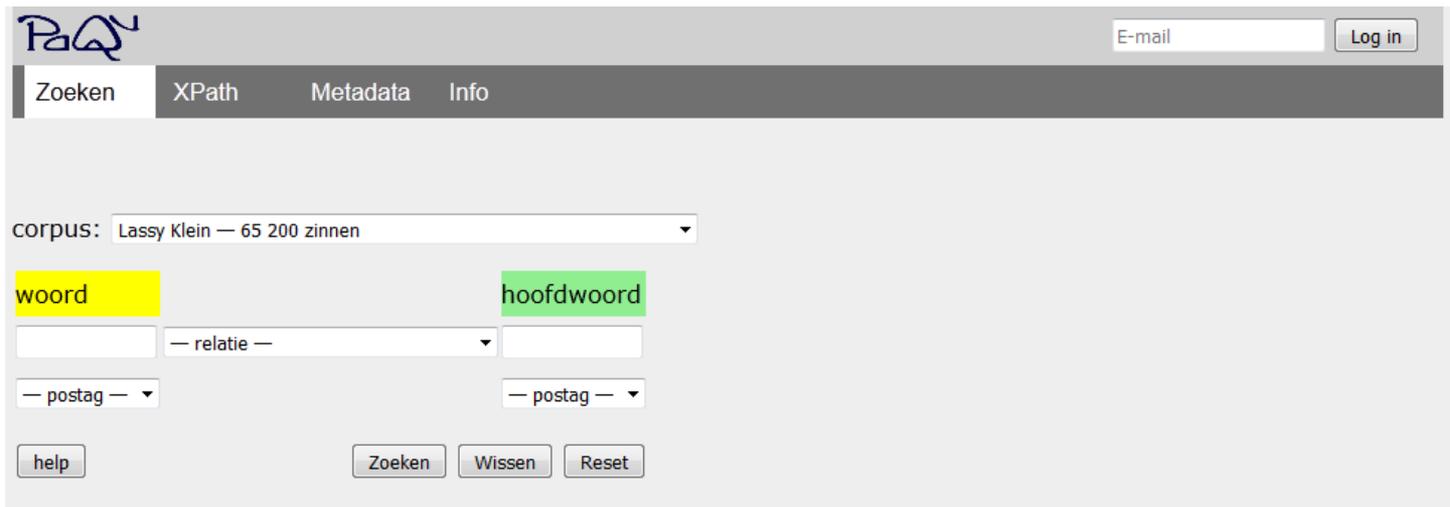
- The (syntactically annotated part of the) Corpus of Spoken Dutch (manually verified syntactic annotation for 1M words of speech)
- The Lassy Small treebank (manually verified syntactic annotation for 1M words of text from various genres)
- The Lassy Large treebank (automatically created syntactic annotation using the Alpino parser for 700M words of text)

# Accessing the corpora

- GrETEL

<http://gretel.ccl.kuleuven.be/gretel-2.0/>

- PAQU <http://zardoz.service.rug.nl:8067/>



The screenshot shows the PAQU web interface. At the top left is the PAQU logo. On the top right, there is an "E-mail" input field and a "Log in" button. Below this is a navigation bar with tabs for "Zoeken", "XPath", "Metadata", and "Info". The "Zoeken" tab is active. The main content area shows a "corpus:" dropdown menu set to "Lassy Klein — 65 200 zinnen". Below this are two input fields: "woord" (highlighted in yellow) and "hoofdwoord" (highlighted in green). Between these fields is a dropdown menu for "relatie" set to "-- relatie --". Below each input field is a dropdown menu for "postag" set to "-- postag --". At the bottom, there are four buttons: "help", "Zoeken", "Wissen", and "Reset".

Example: the possibility of preceding PP-complements:

*Jan was boos over deze opmerking*

Jan was angry over this remark

*Jan was over deze opmerking boos*

```
//node[@cat="ap"]/node[@rel="hd"  
  and @pt="adj"  
  and ../node[@rel="pc"  
  and cat="pp"]/number(@end)=number(@begin)]
```

This query selects the adjectival head of an AP, while the PP-sister of the adjective has to *end* in a position that is equal to the *begin* position of the adjective.

This query has 85 matches in Lassy Small, for 30 different adjectives. However, the adjective *boos* is not among them.

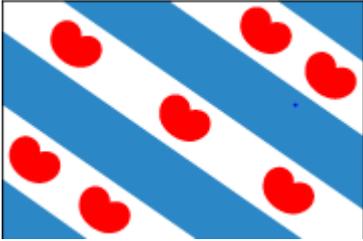
After restricting the query to the lemma *boos*, while removing the linear order constraint, we find that there are 76 hits for *boos*+PP in Lassy Large, but only one for PP+*boos*:

*Leopold II was over die aantasting ... bijzonder boos*

*Leopold II was over that violation ... extremely angry*

This result shows that there is a sharp contrast in frequency between both word orders.

# Open issues



- Morphology and phonology:  
<https://portal.clarin.inl.nl/webcelex/>

# Phonology

- Phonotaxis
  - onsets: hiatus and complex onsets of more than two consonants
  - rhyme: A and B class vowels and diphthongs in open syllables and one consonant coda's
  - Sonority Sequencing Generalization
- phonological processes
  - final devoicing
  - regressive and progressive assimilation
  - degemination

# Morphology

- Inflection:
  - Comparative and superlative
- Derivation:
  - Suffixes: category-changing (*-baar*, *-heid* en *-tje*) and category-neutral (*-schap*)
  - Prefixes: *be-*, *ver-* (category-changing) and *on-* and *oer-* (category-neutral)
  - Non-native prefixes (*pro-* and *sub-*) and suffixes (*-iteit*) and (*-esk*)
- Compounding:
  - NN (*stoelpoot*), VA (*kotsmisselijk*) and ANN (*blotevrouwenblad*)
  - Separable complex verbs (SCVs): *schoonmaken* and *aanvallen*

# Concluding remarks

- Proof of concept: the **Taalportaal** can be enriched with intelligent links to linguistic resources.
- The Taalportaal thus becomes an even more valuable **tool**.
- Linguists are pointed at the existence and usefulness of **search interfaces** developed in the CLARIN infrastructure such as PaQu, GrETEL, OpenSONAR and WEBCELEX.
- By redirecting the user to these front-ends, the **further use** of these applications in the CLARIN infrastructure for modifying queries or submitting new queries is stimulated.