



PAQU DEMO

Jan Odijk

CLARIN 2015 Conference

Wroclaw, 2015-10-16

OVERVIEW

- Queries for *heel, erg, zeer*
 - [CHILDES Adults](#)
 - [Van Kampen Children](#)
- Features
 - [Upload](#)
 - [Search](#)
 - [Analyse](#)
 - [Download / new corpus](#)

UPLOAD



- Upload your own corpus
 - Current formats
 - Plain text (several variants)
 - FoLiA (tokenized)
 - TEI (tokenized)
 - DbXML (DACT)
 - Parsed Alpino files
 - Optionally zipped or tarred



- In
 - LASSY-Small, Spoken Dutch Corpus or LASSY-LARGE/Wikipedia, or
 - Any of your own uploaded corpora
- Search for triples word, relation , head word
 - By lemma, word, root or pos code
 - max 2 can be left unspecified
- Search via XPATH queries



- Analyse search results
 - General statistics per word and headword
 - Combinations
 - Word, relation and head word
 - By lemma, wordform or pos tag
- Analyse metadata (initial version)
 - Currently Spoken Dutch Corpus only
- Analyse search results and metadata
 - Currently Spoken Dutch Corpus only

DOWNLOADS/NEW CORPUS



- Download statistics
- Create New corpus based on a query
 - (limited to 5000 sentences)

ADULTS' UTTERANCES



- Heel in CHILDES Adults
- Erg in CHILDES Adults
- Zeer in CHILDES Adults

- Zeer doen in CHILDES Adults

CHILDREN'S UTTERANCES



- *Heel* in Van Kampen Children
- *Erg* in Van Kampen Children
- *Zeer* in Van Kampen Children

- *Zeer* doen in Van Kampen Children