



Carolin Odebrecht &  
Florian Zipser  
Humboldt-Universität zu Berlin

## **ANNIS workshop**

2014-08-26



- Search and Visualization in Multilayer Linguistic Corpora
  - Imports existing corpora
    - Corpora already have to be annotated, ANNIS only uses what's there
    - No NLP!



- Search and Visualization in Multilayer Linguistic Corpora
  - Makes corpora searchable
    - One query language for all corpora (AQL)
    - Abstraction over linguistic data necessary
    - But: Corpora have different annotations → query has to match the annotations



- Search and Visualization in Multilayer Linguistic Corpora
  - Displays corpora
    - Many visualizations available
    - Corresponding to type of annotation (syntactic trees, phrase trees (RST), grids, coreferences ...)



- What ANNIS cannot do
  - Does not know how to speak natural language  
→ so you have to learn AQL



- What ANNIS cannot do
  - Does not know how to speak natural language  
→ so you have to learn AQL
  - ANNIS does not know any semantics  
→ „NN“, „NP“, „sentence“, „word“, „my favorite annotation“ ... are just sequences of characters



- What ANNIS cannot do
  - Does not know how to speak natural language
    - so you have to learn AQL
  - ANNIS does not know any semantics
    - „NN“, „NP“, „sentence“, „word“, „my favorite annotation“ ... are just sequences of characters
  - You need to be exact
    - e.g. „POS“ != „pos“ and „NN“ != „NN “ (regard the blank)



# ANNIS basics





# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

cat="NP"

Search More History

Status: Ok

Corpus List Search Options

Visible: All

Name	Texts	Tokens		
pcc2	2	399		
RIDGES_Herbology_\	22	122.698		
RIDGES_Herbology_\	29	154.266		

Help/Ex Tutorial Example Queries

Exam	Description	open corpus browser
lemm	search for all words with the lemma "sein" in documents	<a href="#">pcc2</a>
meta:	search for all words with the meta: "Sport"	
exma	search for all Inf-Stat annotations with the value "exmaralda" in the "exmaralda" space	<a href="#">pcc2</a>
"statisch"	search for the word "statisch"	<a href="#">pcc2</a>
lemmal!="sein"	search for all words where the lemma is not "sein"	<a href="#">pcc2</a>
tok!="ist"	search for all words that are not "ist"	<a href="#">pcc2</a>
/.*lich	search for words ending with "lich" (regular expression)	<a href="#">pcc2</a>

Annotations: Enter query, Virtual Keyboard (e.g. arabic), Previous queries, Corpus list



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

cat="NP"

Query Builder

Search

Status: Off

Corpus List

Visible: All

Name	Texts	Tokens		
pcc2	2	399		
RIDGES_Herbology_\	22	122.698		
RIDGES_Herbology_\	29	154.266		

Help/Examples

Tutorial

Example Queries

Example Query	Description	open corpus browser
lemma="sein" & meta::Genre="Sport"	search for all words with the lemma "sein" in documents from the Genre "Sport"	<a href="#">pcc2</a>
exmaralda:Inf-Stat="new"	search for all Inf-Stat annotations with the value "new" in the "exmaralda" namespace	<a href="#">pcc2</a>
"statisch"	search for the word "statisch"	<a href="#">pcc2</a>
lemma!="sein"	search for all words where the lemma is not "sein"	<a href="#">pcc2</a>
tok!="ist"	search for all words that are not "ist"	<a href="#">pcc2</a>
/.*lich/	Search for words ending with "lich" (regular expression)	<a href="#">pcc2</a>

Sample queries (corresponding to corpus)



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

The screenshot displays the ANNIS search interface. On the left, a search box contains 'cat="NP"'. Below it, a 'Query Builder' and 'History' section are visible. A 'Search' button and a 'More' dropdown are present. Below these, it shows '41 matches in 2 documents'. A 'Corpus List' section is visible with a table of corpora.

Name	Texts	Tokens
pcc2	2	399
RIDGES_Herbology_\	22	122.698
RIDGES_Herbology_\	29	154.266
Ridges_Herbology_Ve	13	60.811

The main search results area shows a table of results for the query 'cat="NP"'. The first result is highlighted, showing the path 'pcc2 > 11299 (tokens 1 - 10)'. The table displays the original text and its morphological and syntactic annotations.

Feigenblatt	Die	Jugendlichen	in	Zossen	wollen	ein	M...
Feigenblatt	der	jugendliche	in	Zossen	wollen	ein	M...
Nom.Sg.Neut	Nom.Pl.*	Nom.Pl.*	-	Dat.Sg.Neut	3.Pl.Pres.Ind	Acc.Sg.Neut	Ac...
NN	ART	NN	APPR	NE	VMFIN	ART	NP

Below the table, there are options for visualizations: dependencies (arches), information structure (grid), discourse referents (grid), and constituents (tree). A red box labeled 'Visualizations' points to these options.

The constituents (tree) visualization shows a syntactic tree for the sentence 'Die Jugendlichen in Zossen wollen ein Musikcafé Das'. The root node is 'S', which branches into 'SB', 'HD', and 'OA'. 'SB' branches into 'NP' (circled in red), 'MNR', and 'PP'. 'NP' branches into 'NK' and 'NK'. 'MNR' branches into 'AC' and 'NK'. 'PP' branches into 'AC' and 'NK'. 'HD' branches into 'wollen'. 'OA' branches into 'NP', which further branches into 'NK' and 'NK'. 'S' branches into 'OA', which branches into 'Das'.



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

The screenshot displays the ANNIS interface. On the left, a search results table shows 41 matches in 2 documents. The 'Corpus List' is visible, with 'pcc2' selected. A red box highlights the 'Corpus metadata window' on the right, which provides detailed information for the selected corpus.

**Corpus metadata window (pcc2, ID: 5664):**

Name	Value
URL	<a href="#">link</a>
annotation_description	POS, lemma, morphology, constituent and dependency syntax, information structure, coreference, rhetorical structure, article headings
annotation_levels	pos;lemma;morph;Inf-Stat;Focus_newInf;PP;NP;Topic;Sent;Foc_c (for dominance egdes);dep:func (for dependency pointing relations);anaphor_antecedent (pointing relations)
full_name	Potsdam Commentary Corpus (sample of 2 documents)
language	German
source	Project D1, SFB 632
version	6.0

Name	Example (click to use query)	URL
exmaralda:Focus	exmaralda:Focus_newInf="nf-unsol"	<a href="#">↔</a>
exmaralda:headi	exmaralda:heading="heading"	<a href="#">↔</a>
exmaralda:Inf-St	exmaralda:Inf-Stat="giv-active"	<a href="#">↔</a>
exmaralda:NP	exmaralda:NP="NP"	<a href="#">↔</a>
exmaralda:PP	exmaralda:PP="PP"	<a href="#">↔</a>
exmaralda:Sent	exmaralda:Sent="s"	<a href="#">↔</a>
exmaralda:Topic	exmaralda:Topic="ab"	<a href="#">↔</a>
mmax:ambiguity	mmax:ambiguity="not_ambig"	<a href="#">↔</a>
mmax:anaphor_i	mmax:anaphor_type="anaphor_nomin	<a href="#">↔</a>
mmax:complex_	mmax:complex_np="no"	<a href="#">↔</a>
mmax:dir_speec	mmax:dir_speech="text_level"	<a href="#">↔</a>
mmax:grammati	mmax:grammatical_role="other"	<a href="#">↔</a>
mmax:np_form	mmax:np_form="defnp"	<a href="#">↔</a>
mmax:phrase_ty	mmax:phrase_type="np"	<a href="#">↔</a>
mmax:referentia	mmax:referentiality="discourse-new"	<a href="#">↔</a>
mmax:type	mmax:type="none"	<a href="#">↔</a>
rst:kind	rst:kind="segment"	<a href="#">↔</a>

Link to corpus: [https://korpling.german.hu-berlin.de/annis3/#\\_c=cGNjMg](https://korpling.german.hu-berlin.de/annis3/#_c=cGNjMg)



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

Document metadata

Document metadata window

1 Path: pcc2 > 11299 (tokens 1 - 10)

Feigenblatt	Die	Jugendlichen	in	Zossen	wollen	ein
Feigenblatt	der	jugendliche	in	Zossen	wollen	ein
Nom.Sg.Neut						Acc.Sg.Neut
NN						ART

Info for salt:/pcc2/11299

### Metadata

document: 11299

Name	Value
Dokumentname	pcc-11299
Genre	Politik
Titel	Feigenblatt

corpus: pcc2

dependencies (arch...)  
information structur...  
discourse referents...  
constituents (tree)



- Basic principles of AQL (ANNIS Query Language)
  - Attributes and values
    - Searching for exact character sequences
    - Searching for patterns
  - Combinatory search



- Corpus for demonstration: pcc2 (a sub corpus of pcc)

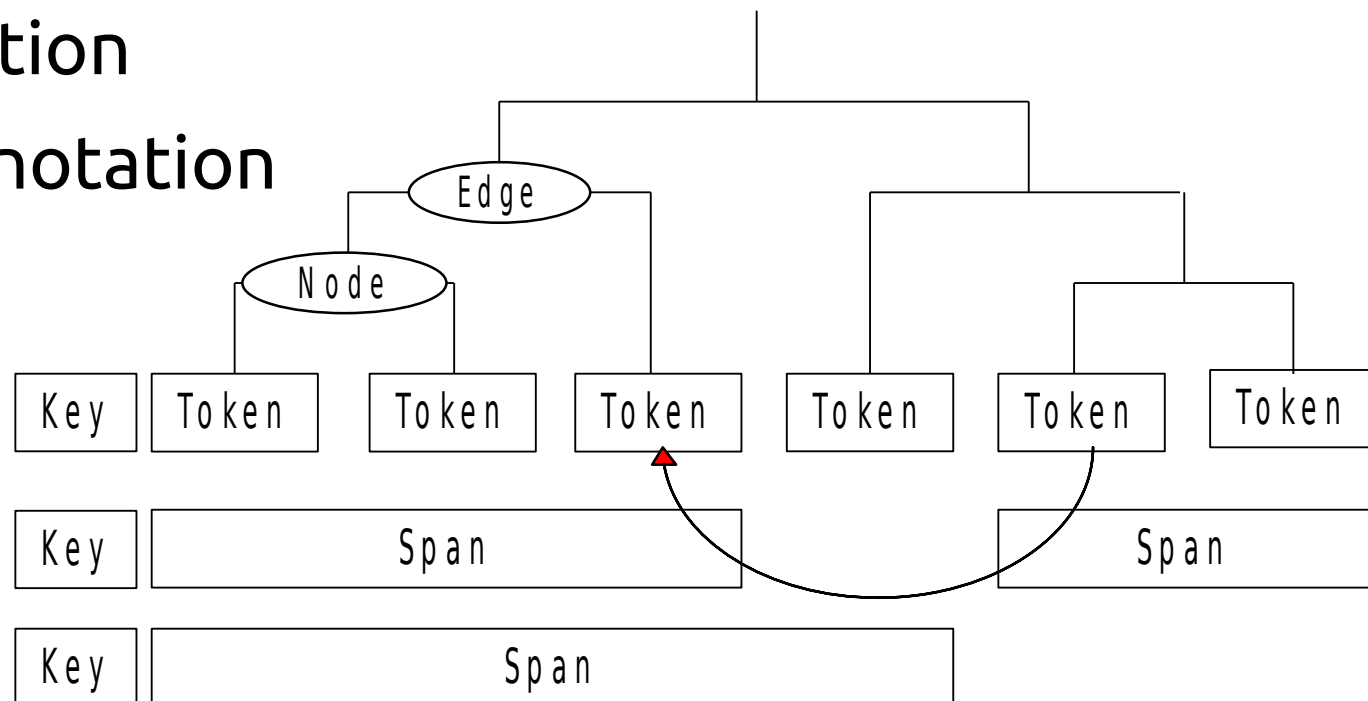
[https://korpling.german.hu-berlin.de/annis3/#\\_c=cGNjMg](https://korpling.german.hu-berlin.de/annis3/#_c=cGNjMg)

- Potsdam Commentary Corpus
  - German Newspaper commentaries  
'Märkische Allgemeine Zeitung'

<https://www.ling.uni-potsdam.de/acl-lab/Forsch/pcc/pcc.html>

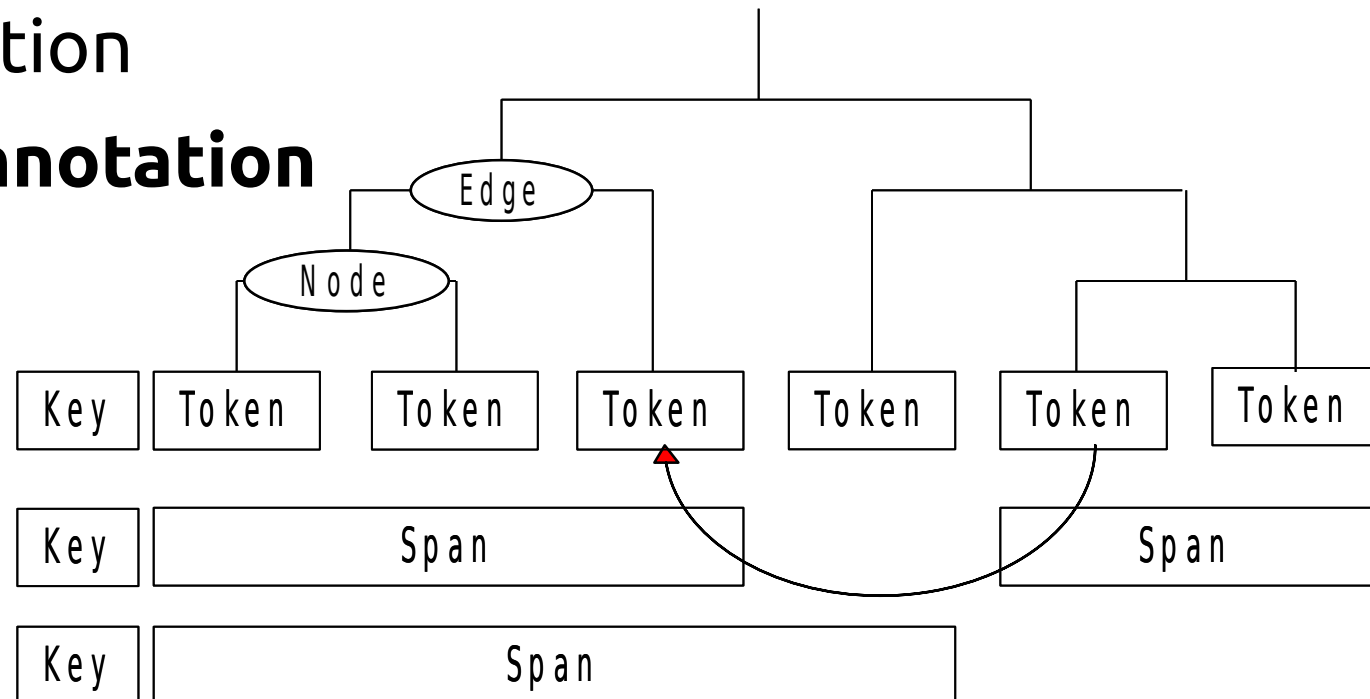
- Multiple annotations

- Different types of annotations
  - Token annotation
  - Span annotation
  - Pointing relation
  - Hierarchy annotation (trees)





- Different types of annotations
  - **Token annotation**
  - **Span annotation**
  - Pointing relation
  - **Hierarchy annotation (trees)**





- Token annotation
  - Exact sequence

searching for a word form

"Jugendlichen"

"jugendlichen"

- Token annotation
  - Exact sequence

searching for a word form

"Jugendlichen" 3 hits

"jugendlichen" 0 hits

→ tok="jugendlichen"

- Token annotation
  - Exact sequence

searching for an **exact** part of speech tag

pos	=	"NN"
-----	---	------

attribute                      value

- Attributes can have more than one value
- Searching for all values of an attribute



- Token annotation
  - Exact sequence

searching for an exact part of speech tag

`pos="NN"`

`pos="ADJA"`

- Token annotation
  - Exact sequence

searching for an exact part of speech tag

`pos="NN"` 62 hits

`pos="ADJA"` 18 hits

searching for all values of an attribute

`pos` 399 hits



- Span annotation
  - Exact sequence

searching for sentences

Sent="s"



- Span annotation
  - Exact sequence

searching for sentences

`Sent="s"` 28 hits





- **Sent="s"** 28 hits
  - necessary to know which annotations are in a corpus

Visible: All

Name	Texts	Tokens
pcc		
pcc2	2	399

Corpus information for pcc2 (ID: 5664)

Metadata		Available annotations	
Select corpus/document:	pcc2	Node Annotations	
Name	Value	Name	Example (click to use query)   URL
URL	<a href="#">link</a>	exmaralda:Focu	exmaralda:Focus_newInf="nf-unsol"
annotation_description	POS, lemma, morphology, constituent and dependency syntax, information structure, coreference, rhetorical structure, article headings	exmaralda:head	exmaralda:heading="heading"
annotation_levels	pos;lemma;morph;Inf-Stat;Focus_newInf;PP;NP;Topic;Sent;Foc_c (for dominance egdes);dep:func (for dependency pointing relations);anaphor_antecedent (pointing relations)	exmaralda:Inf-S	exmaralda:Inf-Stat="giv-active"
full_name	Potsdam Commentary Corpus (sample of 2 documents)	exmaralda:NP	exmaralda:NP="NP"
language	German	exmaralda:PP	exmaralda:PP="PP"
source	Project D1, SFB 632	exmaralda:Sent	exmaralda:Sent="s"
version	6.0	exmaralda:Topic	exmaralda:Topic="ab"
		mmax:ambiguity	mmax:ambiguity="not_ambig"
		mmax:anaphor_	mmax:anaphor_type="anaphor_nomi
		mmax:complex_	mmax:complex_np="no"
		mmax:dir_speec	mmax:dir_speech="text_level"
		mmax:grammati	mmax:grammatical_role="other"
		mmax:np_form	mmax:np_form="defnp"
		mmax:phrase_t	mmax:phrase_type="np"
		mmax:referentie	mmax:referentiality="discourse-new"
		mmax:type	mmax:type="none"
		rst:kind	rst:kind="segment"
		Edge Annotations	
		Edge Types	
		Meta Annotations	

Link to corpus: [https://korpling.german.hu-berlin.de/annis3/#\\_c=cGNjMg](https://korpling.german.hu-berlin.de/annis3/#_c=cGNjMg)



- Token annotation
    - Patterns
      - matches any single character
      - \* zero or more of the preceding element
- searching for the beginning a of word

```
/Jugend.*/
```

```
/jugend.*/
```

- Token annotation

- Patterns

- matches any single character

- \* zero or more of the preceding element

searching for the beginning a of word

`/Jugend.*/`

5 hits ("Jugendlichen" 3 hits)

*Jugendlichen Jugendliche*

`/jugend.*/`

0 hits ("jugendlichen" 0 hits)



- Token annotation
  - patterns

searching for **all** nouns

`pos=/N./`

includes NN & NE

searching for **all** adjectives

`pos=/ADJ./`

includes ADJA & ADJD



- Token annotation
  - patterns

searching for **all** nouns

`pos=/N./`

73 hits (pos="NN" 62 hits)

searching for **all** adjectives

`pos=/ADJ./`

32 hits (pos="ADJA" 18 hits)

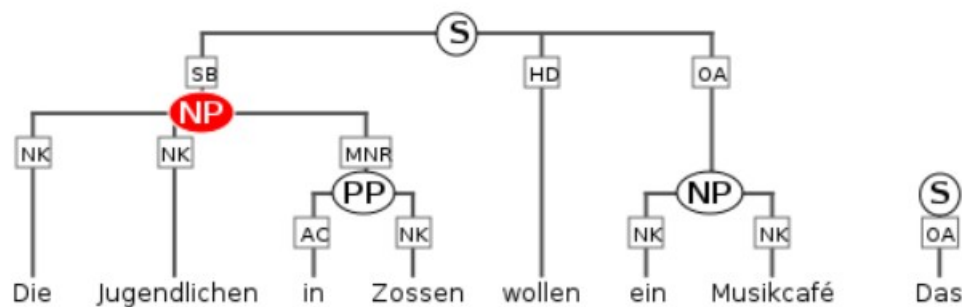
- Span annotation

searching for all NPs

`cat="NP"`

41 hits (pos="NN" 62 hits)

e.g. *Die Jugendlichen in Zossen*



- Relations between attributes

searching for all NPs which contain a preposition

`cat="NP"`

41 hits

`pos="APPR"`

19 hits

e.g. *Die Jugendlichen in Zossen*

→ no relation between the two information!

- Relations between attributes

searching for **all** NPs which contain a preposition

`cat="NP"`

#1

`pos="APPR"`

#2

e.g. *Die Jugendlichen in Zossen*

→ NP includes APPR



- Relations between attributes

searching for **all** NPs which contain a preposition

```
cat="NP" &
pos="APPR" &
#1_i_#2
```

e.g. *Die Jugendlichen in Zossen*

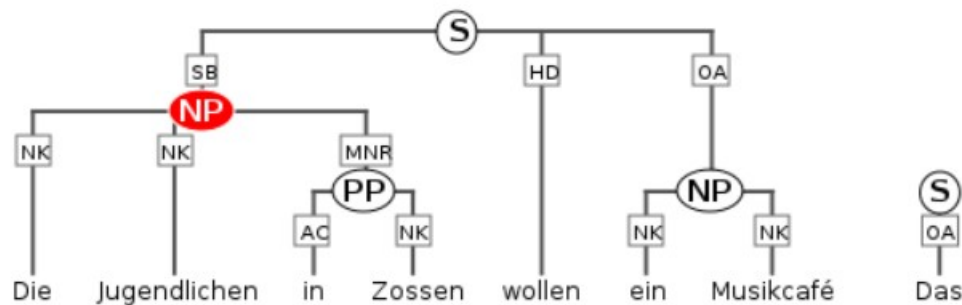


- Relations between attributes

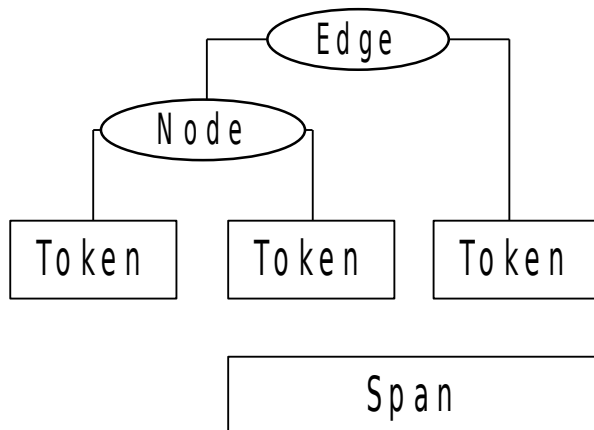
searching for **all** NPs which are **objects**

`cat="NP"`

e.g. *Die Jugendlichen in Zossen* -->subject!



- Relations between attributes
  - searching all **NPs** which are **objects**
  - **NP** → node annotation
  - **OA** → edge annotation



- Relations between attributes

searching all NPs which are objects

```
cat="NP"
```

the syntactic function in the tree

```
func="OA"
```

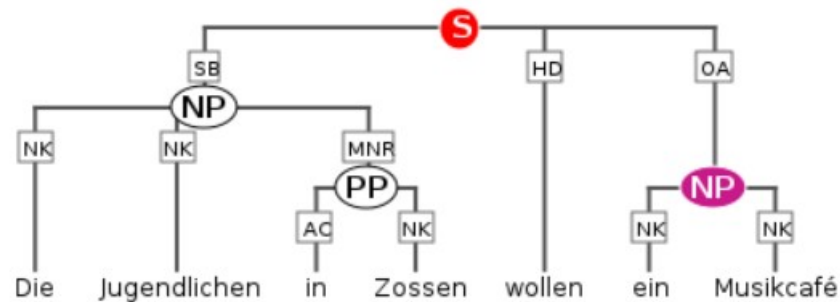
→ Note: At least there are two elements which relate in a way to each other!

- Relations between attributes

searching all NPs which are objects

`node & cat="NP" & #1 >[func="OA"] #2`

e.g. *ein Musikcafé* --> object!



- Relations we used:

A \_i\_ B

A includes B

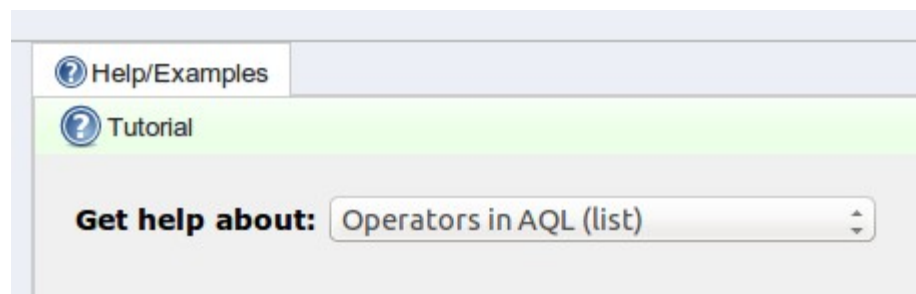
A > B

A dominates B

A >[func="OA"] B

A dominates B and B is an  
object

The full list of relations can be found in ANNIS





ANNIS: Search and  
Visualization in  
Multilayer Linguistic  
Corpora

What's new in  
ANNIS

# What's new in ANNIS

## version 3.1.7

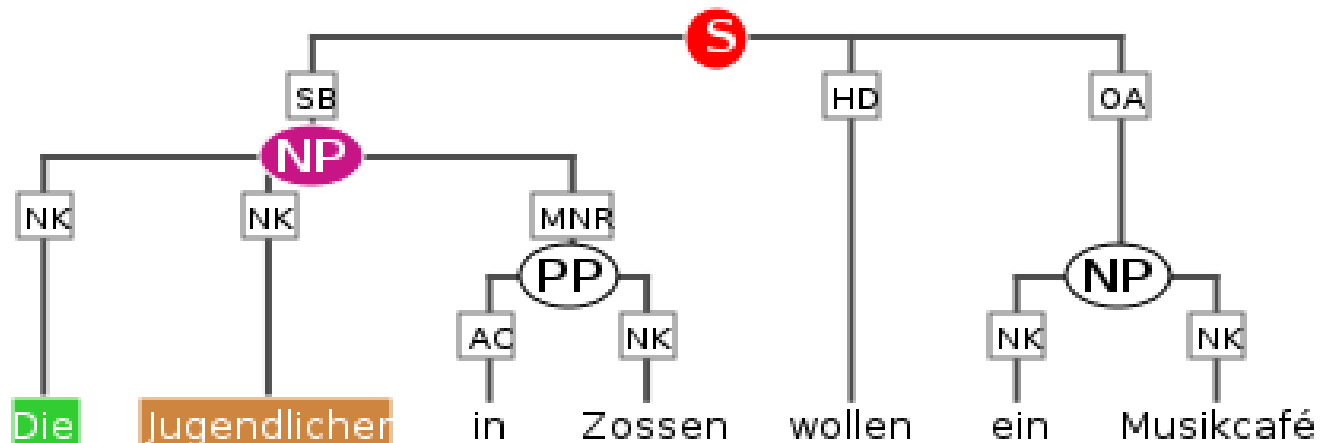


- Simplified syntax (AQL)
- Frequency analysis (Visualisierung)
- Expand match context (Visualisierung)
- Equality and Inequality (AQL)
- Variables (AQL)
- Complex OR expression (AQL)
- Document browser (Visualisierung)
- CSV export (Visualisierung)
- Tooltip for corpus names (Visualisierung)
- Report problem (Visualisierung)



- Question:

„Die“ followed by „Jugendlichen“ both being dominated by a prepositional phrase which is dominated by a sentence

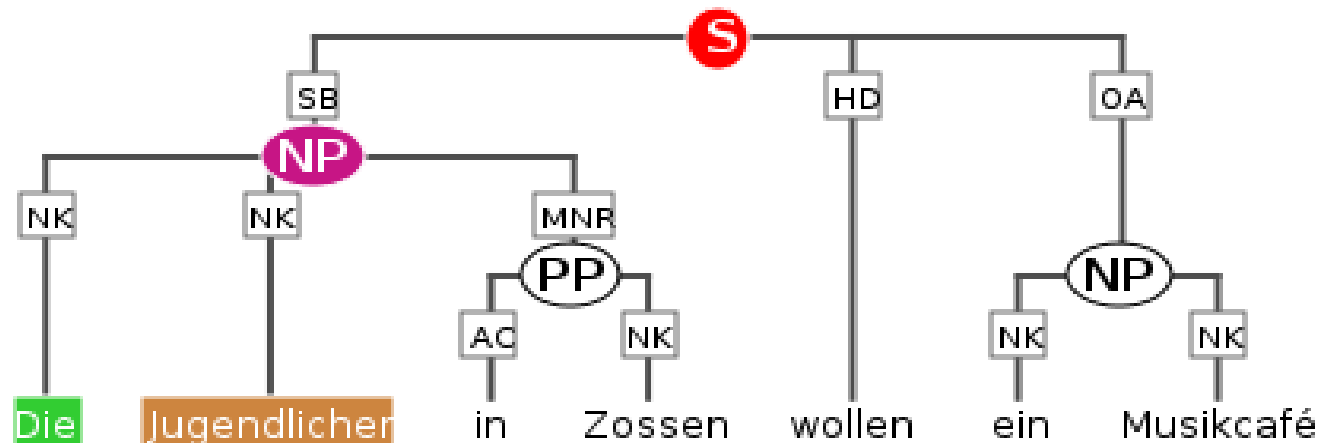


So far:

cat="S" & cat="NP" & "Die" & "Jugendlichen" & #1 > #2 & #2 > #3 & #2 > #4 & #3 . #4

- Question:

„Die“ followed by „Jugendlichen“ both being dominated by a prepositional phrase which is dominated by a sentence



So far:

cat="S" & cat="NP" & "Die" & "Jugendlichen" & #1 > #2 & #2 > #3 & #2 > #4 & #3 . #4

Simplified:

cat="S" > cat="NP" > "Die" . "Jugendlichen" & #2 > #4



- Question:
  - How many words tagged as „NN“, „ADJA“ or „ADV“ does a corpus contain?
  - What are the most frequent part-of-speech tags followed by a noun?
  - What are the most frequent part-of-speech tags in a prepositional phrase, which is in a sentence?
  - ...



The screenshot shows the ANNIS interface with a search query 'pos' entered. The search results show 399 matches in 2 documents. The 'More' dropdown menu is open, highlighting the 'Frequency Analysis' option. The 'Frequency Analysis' window is also visible, showing the selected corpora 'pcc2' and the query 'pos'. The table below shows the results of the frequency analysis.

Node nun	Selected annotation of	Comment
1	pos	automatically created from pos

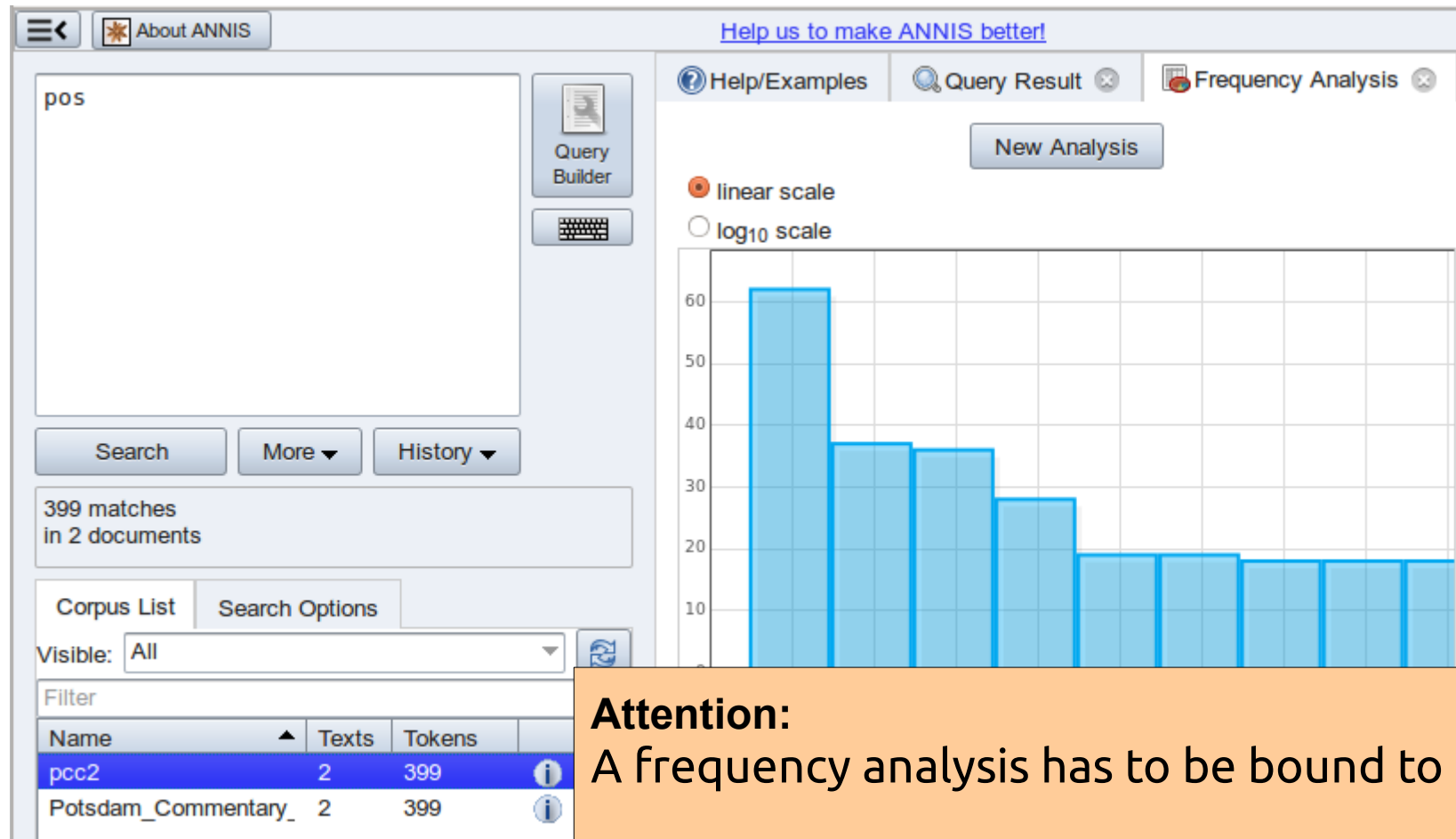
Buttons for 'Export' and 'Frequency Analysis' are highlighted with orange arrows. The 'Perform frequency analysis' button is also highlighted with an orange arrow.



The screenshot shows the ANNIS interface with a search for 'pos' resulting in 399 matches in 2 documents. The frequency analysis bar chart displays the following data:

POS Tag	Count
NW	62
ADV	37
ART	36
\$.	28
\$.	19
APPR	19
ADJA	18
PPER	18
VVFIN	18

An orange arrow points to the 'Download as CSV' button at the bottom right of the chart area.



The screenshot shows the ANNIS interface with a search for 'pos' resulting in 399 matches in 2 documents. A frequency analysis bar chart is displayed on the right, showing the distribution of results. The chart has a linear scale selected. The y-axis ranges from 0 to 60, and the x-axis shows 10 bins. The first bin has the highest frequency, around 62, and subsequent bins show a general downward trend.

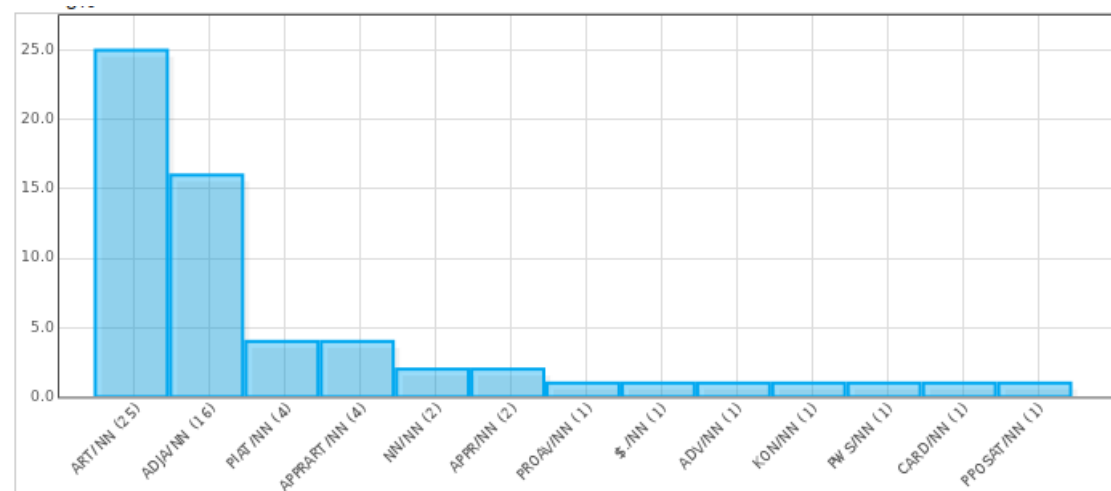
**Attention:**  
A frequency analysis has to be bound to a query!

Name	Texts	Tokens
pcc2	2	399
Potsdam_Commentary_	2	399



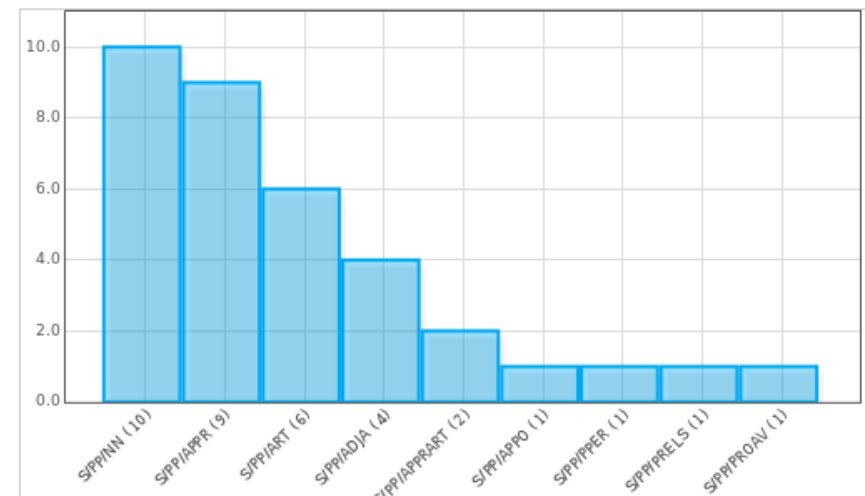
- What are the most frequent part-of-speech tags followed by a noun?

```
pos . pos="NN"
```

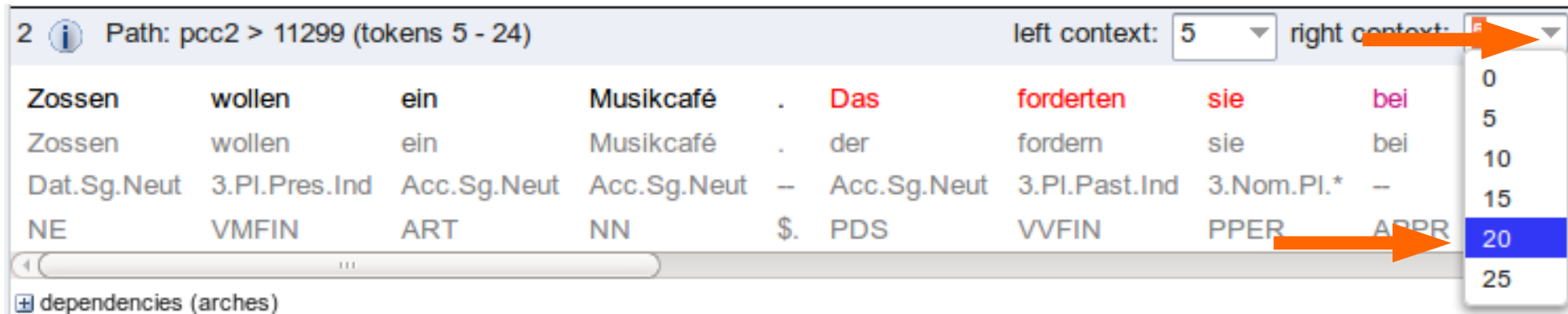


- What are the most frequent part-of-speech tags in a prepositional phrase, which is in a sentence?

```
cat="S" > cat="PP" > pos
```



- Sometimes the context is too small



2 Path: pcc2 > 11299 (tokens 5 - 24) left context: 5 right context: 5

Zossen	wollen	ein	Musikcafé	.	Das	forderten	sie	bei
Zossen	wollen	ein	Musikcafé	.	der	fordern	sie	bei
Dat.Sg.Neut	3.PI.Pres.Ind	Acc.Sg.Neut	Acc.Sg.Neut	-	Acc.Sg.Neut	3.PI.Past.Ind	3.Nom.PI.*	-
NE	VMFIN	ART	NN	\$.	PDS	VVFIN	PPER	APPR

dependencies (arches)

- Even more than 25 is possible, it's a free text field





- Equality „==“ and inequality „!=“ for attributes
- Question (equality):  
two same part-of-speech tags, one directly following the other

ersten

Zossener

erster

Zossener

Pos.Dat.Sg.Fem

Pos.\*.\*

ADJA

ADJA

pos . pos & #1 == #2

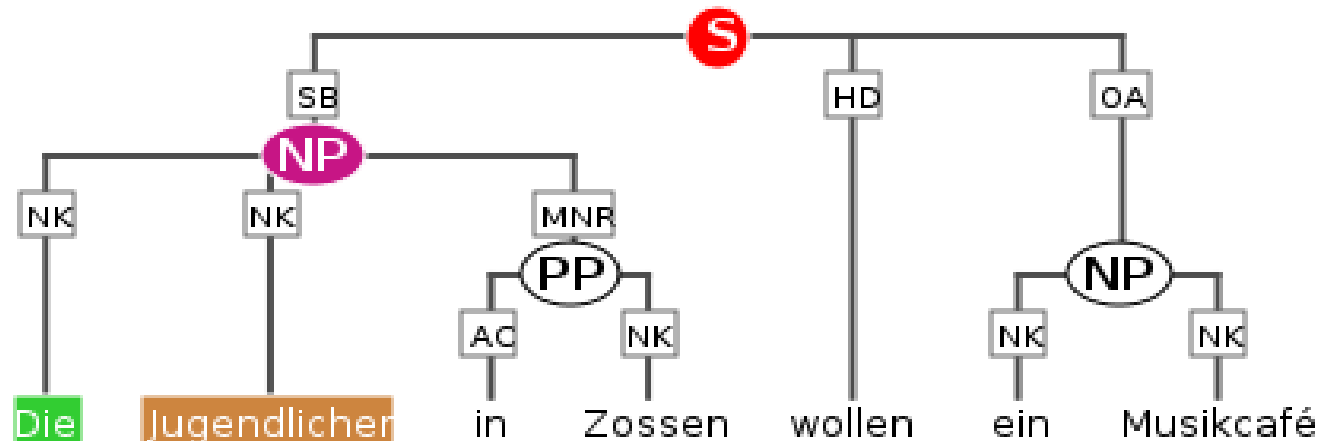
- Equality „==“ and inequality „!=“ for attributes
- Question (inequality):  
two different part-of-speech tags, one directly following  
the other

Die	Jugendlichen
der	jugendliche
Nom.Pl.*	Nom.Pl.*
ART	NN

pos . pos & #1 != #2

- Question:

„Die“ followed by „Jugendlichen“ both being dominated by a prepositional phrase which is dominated by a sentence

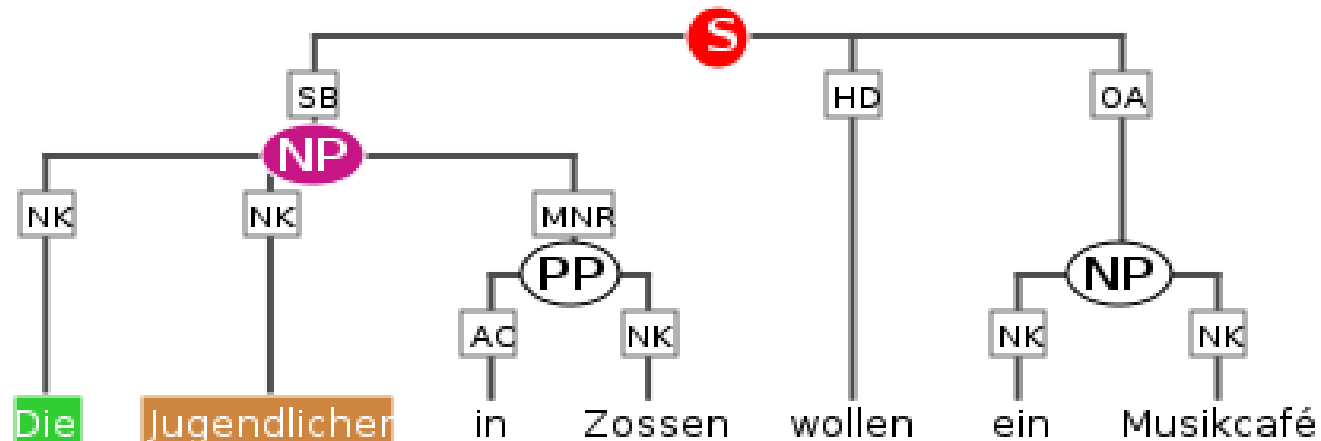


Simplified:

`cat="S" > cat="NP" > "Die" . "Jugendlichen" & #2 > #4`

- Question:

„Die“ followed by „Jugendlichen“ both being dominated by a prepositional phrase which is dominated by a sentence

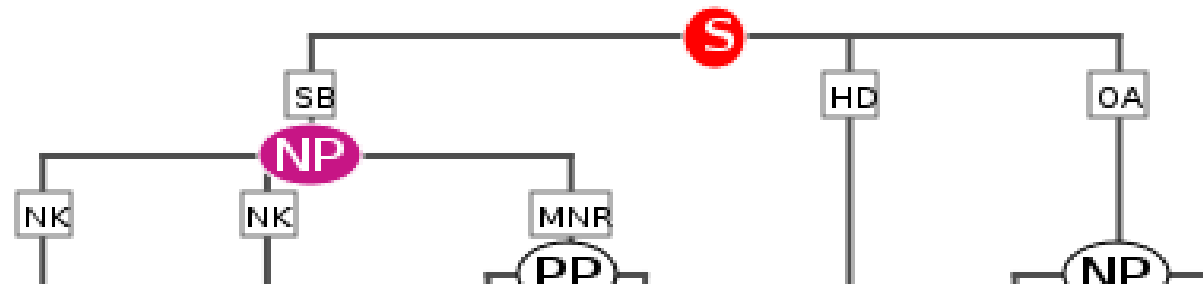


Simplified:

`cat="S" > np#cat="NP" > "Die" . jug#"Jugendlichen" & #np > #jug`

- Question:

„Die“ followed by „Jugendlichen“ both being dominated by a prepositional phrase which is dominated by a sentence



**Variables and numbers can be mixed:**

`cat="S" > np#cat="NP" > "Die" . "Jugendlichen" & #np > #4`

Simplified:

`cat="S" > np#cat="NP" > "Die" . jug#"Jugendlichen" & #np > #jug`



- Question (simple OR):

A part-of-speech tag which is a noun, an attributive adjective or an article

```
pos=/(NN)|(ADJA)|(ART)/      (in pattern search)
```



- Question (simple OR):

A part-of-speech tag which is a noun, an attributive adjective or an article

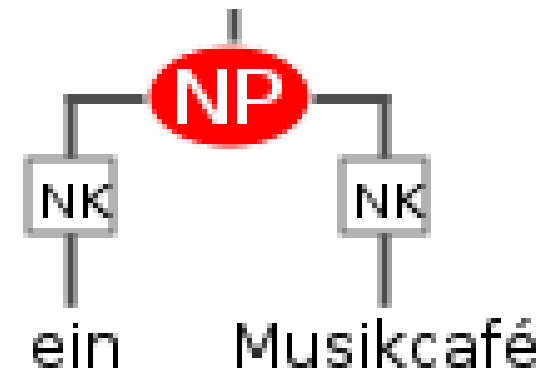
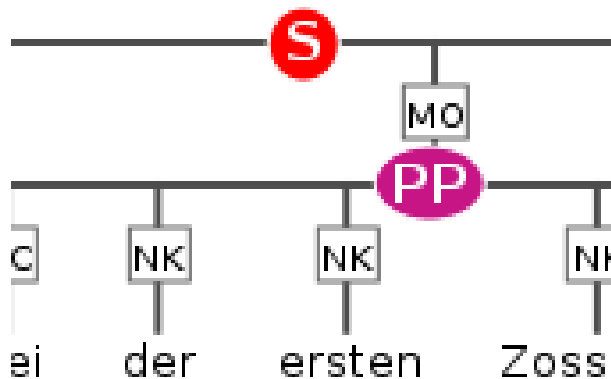
```
pos=/(NN)|(ADJA)|(ART)/      (in pattern search)
```

- OR for expressions

```
pos="NN" | pos="ADJA" | pos="ART"
```

- Question (complex OR):

A prepositional phrase, which is dominated by a sentence, or just a nominal phrase

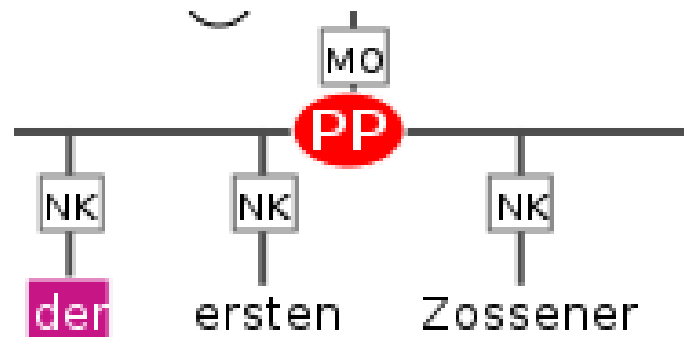


`(cat="S" > cat="PP") | cat="NP"`



- Question (nested OR):

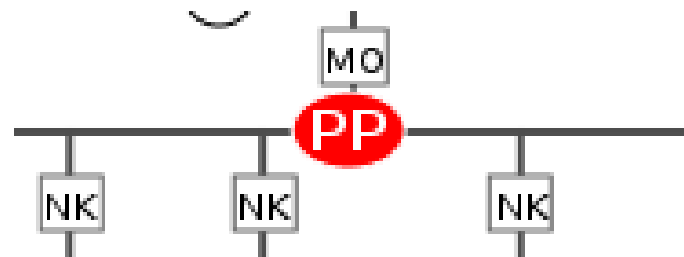
A prepositional phrase, which dominates a noun, an attributive adjective or an article



`a#cat="PP" &`  
`(b#pos="NN" | b#pos="ADJA" | b#pos="ART") &`  
`#a > #b`

- Question (nested OR):

A prepositional phrase, which dominates a noun, an attributive adjective or an article



**Attention:**  
 All expressions in brackets have to use the same variable  
 ... & (b#pos="NN" | b#pos="ADJA" | b#pos="ART") & ...

a#cat="P

(b#pos="NN" | b#pos="ADJA" | b#pos="ART") &

#a > #b



- Displays the entire text of a document

The screenshot shows the ANNIS web interface. On the left, there is a search area with a text input field containing "Please enter AQL query", a "Query Builder" button, and a keyboard icon. Below this are "Search", "More", and "History" buttons. A status bar shows "Status: Ok". At the bottom left, there is a "Corpus List" section with a "Visible: All" dropdown and a "Filter" input field. A table below the filter shows the following data:

Name	Texts	Tokens
pcc2	2	399

On the right side of the interface, there is a "Help us to make ANNIS better!" link and a "Help/Examples" button. Below these is a table with the following data:

document name	corpus path	visualizer	info
11299	pcc2 > 11299	<a href="#">full text</a>	
4282	pcc2 > 4282	<a href="#">full text</a>	

An orange arrow points to the "full text" link in the first row of the table. Another orange arrow points to the "Visible: All" dropdown in the "Corpus List" section.



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

# Document browser

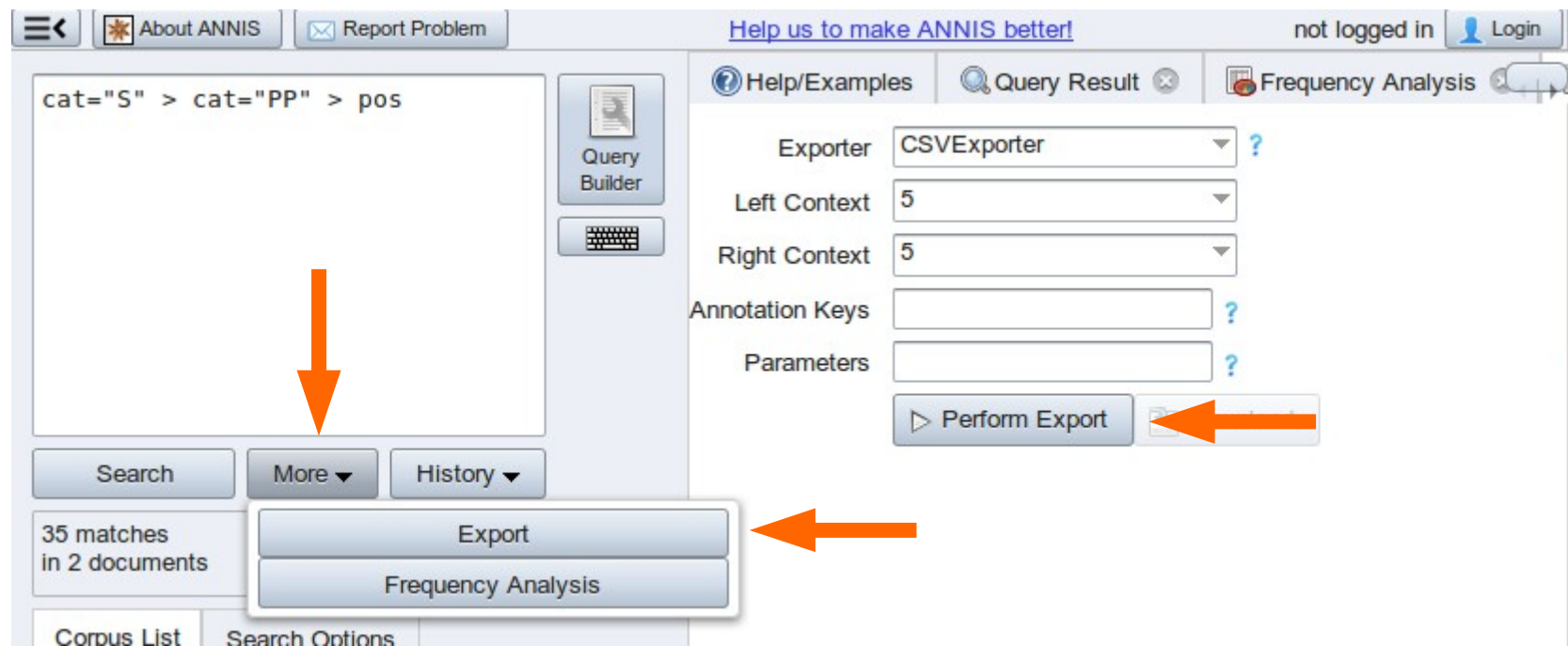
The screenshot shows the ANNIS web interface. On the left, there is a search input field with the placeholder text "Please enter AQL query". Below it are buttons for "Search", "More", and "History". A "Query Builder" button is also visible. The status bar shows "Status: Ok". Below that, there are tabs for "Corpus List" and "Search Options". The "Visible:" dropdown is set to "All". A table lists the corpora:

Name	Texts	Tokens
pcc2	2	399

On the right, there is a browser window titled "Help us to make ANNIS better!". The active tab is "pcc2 > 11299 - Visualizer: full text". The main content area displays a text snippet:

Feigenblatt Die Jugendlichen in Zossen wollen ein Musikcafé . Das forderten sie bei der ersten Zossener Runde am Dienstagabend . Dass die Politiker der Stadt dafür Verständnis haben , ist löblich . Mit dem Treffen im Rathaus ist somit auch ein Dialog zwischen den Generationen angestoßen . Dass die beiden geladenen Jugendlichen im Laufe des Abends immer weniger zu Wort kamen , war sicher keine böse Absicht , ärgerlich ist es trotzdem . Und aberwitzig dazu . Es entbehrt nicht der Komik , wenn sich drei Erwachsene - Karola Andrae ( Bürgerbündnis/FDP ) , Susanne Michler ( CDU ) und Joachim Zanow ( SPD ) - darüber streiten , was Jugendliche wollen und brauchen , ohne auf die Idee zu kommen , sie selbst zu fragen . Und das , obwohl sie ihnen gegenüber sitzen . Die Jugendlichen wurden somit zum bloßen Feigenblatt degradiert . Nicht über sondern mit ihnen hätten die Politiker reden sollen . Damit ist eine große Chance vertan . Vielleicht klappt es bei der nächsten Runde Anfang 2002 . Dann werden auch mehr Jugendliche eingeladen . In der Gruppe können sie sich hoffentlich mehr Gehör verschaffen . Vielleicht finden dann auch Vertreter von PDS und Gewerbeverein ihren Weg ins Rathaus . Die glänzten diesmal noch mit Abwesenheit .


- Export data for further processing







The screenshot shows the ANNIS web interface. At the top, there are navigation links: 'About ANNIS', 'Report Problem', and 'Help us to make ANNIS better!'. The user is 'not logged in' and there is a 'Login' link. The search query is 'cat="S" > cat="PP" > pos'. The search results show '35 matches in 2 documents'. The 'More' dropdown menu is open, showing 'Export' and 'Frequency Analysis' options. The 'Export' option is highlighted with an orange arrow. The 'Perform Export' button is also highlighted with an orange arrow. The interface includes a search bar, a query builder, and a results section showing 35 matches in 2 documents.

- Sometimes corpus names can get very long

Corpus List    Search Options

Visible: All 

Filter

Name ▲	Texts	Tokens		
pcc2	2	399		
Potsdam_Commenta	2	399		

Potsdam\_Commentary\_Corpus



# ANNIS: Search and Visualization in Multilayer Linguistic Corpora

# Report problem

The screenshot shows the 'Report Problem' form in the ANNIS interface. The form includes the following fields:

- Short Summary\***: A text input field.
- Long Description\***: A text area with the prompt "What steps will reproduce the problem?" followed by a numbered list (1., 2., 3.) and the prompt "What is the expected result?". Below this is another text area with the prompt "What happens instead".
- Your Name\***: A text input field.
- Your e-mail address\***: A text input field.

Below the form is a link labeled "Hide attached screenshot". An orange arrow points to the "Report Problem" button in the top navigation bar. Another orange arrow points to the "Submit problem report" button at the bottom of the attached screenshot.

The attached screenshot shows the ANNIS search interface with a list of search results. The results are displayed in a table with columns for document name, frequency, and other metadata. The table includes entries such as "FalkoSummary\_1v12", "FalkoSummary\_2v12", "FalkoSummary\_3v12", etc. The "Submit problem report" button is highlighted with an orange arrow.



- ANNIS comes in two flavors
  - A server version
  - A desktop version (ANNIS kickstarter)
  - Both are downloadable at:  
<http://www.sfb632.uni-potsdam.de/annis/>
- ANNIS is open source (Apache license 2.0) and hosted on github
  - <https://github.com/korpling/ANNIS>





Thanks for your attention!  
Any questions?

[carolin.odebrecht@hu-berlin.de](mailto:carolin.odebrecht@hu-berlin.de),  
[f.zipser@gmx.de](mailto:f.zipser@gmx.de)