

## "Pretty": adjective vs. adverb (BNC + BNCweb)

- *Pretty* can occur as an adjective: "She is a pretty woman."
- *Pretty* can occur as an adverb: "She is a pretty ugly woman."

### Task:

- *Pretty* as an adjective can premodify nouns or be used in predicative function. *Pretty* as an adverb occurs mostly before adjectives and other adverbs.
  - The BNC is a Part-of-Speech-tagged corpus in which the disambiguation has been carried out by a semi-automated tagging procedure. Use the Part-of-Speech (POS) tags to find out frequencies of these grammatical functions in the different sections of the BNC.
- Compare the relative percentages of *pretty* (Adj) and *pretty* (Adv) in the different subdivisions of the corpus.

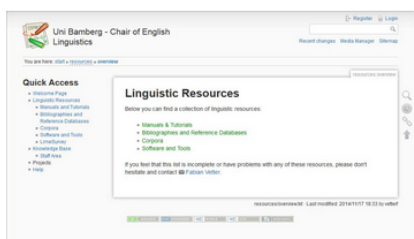
### Use the BNCweb corpus:

## Available Web Services & Linguistic Resources

Chair of English Linguistics @ University of Bamberg

### Welcome!

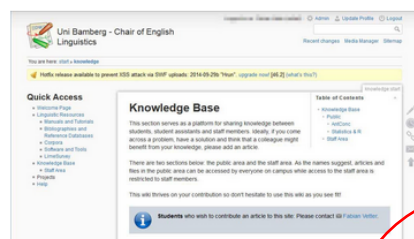
Below you will find an overview of our web services. Naturally, all of them can be used **free of charge**. You can [sign up for a free user account here](#). Your account grants you access to the knowledge base, the corpora (BNCweb, COCA, COHA, GloWbE (coming soon)), RStudio Server. Please note: most services are only available from within the university's network. If you require access to LimeSurvey, please contact [Fabian Vetter](#).



### (Corpus-) Linguistic Resources

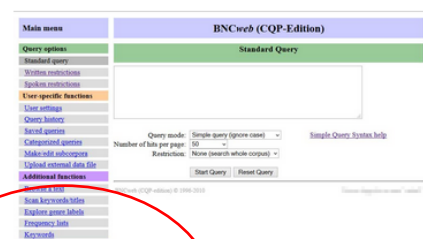
Here you can find a collection of linguistic resources:

- information and links to databases
- links to linguistic programs
- manuals & tutorials
- fonts for phonetic transcriptions
- ...



### Knowledge Base

A simple knowledge base for staff and students alike. Provides a public section and a staff area.  
**Registration required.**  
**VPN connection recommended.**  
-> [Registration](#).



### BNCweb

Web interface to the British National Corpus.  
**Registration & VPN connection required.**  
-> [Registration](#).

https://eng-ling.uni-bamberg.de/bncweb-cgi/BNCQuery.pl?theQuery=search&urlTest=yes

**Main menu** **BNCweb (CQP-Edition)**

**Query options** **Standard Query**

Standard query

[Written restrictions](#)

[Spoken restrictions](#)

**User-specific functions**

[User settings](#)

[Query history](#)

[Saved queries](#)

[Categorized queries](#)

[Make/edit subcorpora](#)

[Upload external data file](#)

**Additional functions**

[Browse a text](#)

[Scan keywords/titles](#)

[Explore genre labels](#)

[Frequency lists](#)

[Keywords](#)

**About BNCweb**

[New: BNCweb book](#)

[The BNCweb team](#)

[New features](#)

[Bug reports](#)

[The CLAWS-5 tagset](#)

[Oxford BNC homepage](#)

Insert the word you want to search including the POS tags

Query mode: Simple query (ignore case) [Simple Query Syntax help](#)

Number of hits per page: 50

Restriction: None (search whole corpus)

BNCweb (CQP-edition) © 1996-2010

Version 4.3, January 2010

### Matching parts-of-speech (POS)

- Search for a word form with a specific POS tag by linking them with an underscore `_`. Wildcards can be used both for word form and POS tag:

<code>lights_NN2</code>	→ plural noun <i>lights</i> , but not the verb form <i>lights</i>
<code>*ly_AJ0</code>	→ adjectives ending in <i>-ly</i> (e.g. <i>daily</i> )
<code>super+_V*</code>	→ verb forms starting with <i>super-</i>

- You can also search by POS tag only: `_PNX` → any reflexive pronoun
- Complete listing of POS tags used in the BNC is given on last page.
- Use simplified POS tags enclosed in curly braces: `super+_{VERB}` for verb forms starting with *super-* (no wildcards allowed in simplified tags).
- List of simplified POS tags (Table 3.8 of Hoffmann *et al.* (2008) shows comparison with full tagset):

A, ADJ	adjective	INT, INTERJ	interjection
N, SUBST	noun	PREP	preposition
V, VERB	verb	PRON	pronoun
ADV	adverb	\$, STOP	punctuation
ART	article	UNC	other / uncertain
CONJ	conjunction		

- Keep in mind that part-of-speech tags have been assigned by an automatic software tool and are not always correct (try e.g. `beer_{N}` `can_{N}`).

What you need to know about part-of-speech (POS) tagging:

## Evaluating the results:

### - Standard query: pretty\_AJO (Adjective)

Your query "pretty\_AJO" returned 2613 hits in 891 different texts. The current solution set was found in 891 texts. [View results](#)

Categories:

Categories (for crosstabs only):

**The following distribution was found:**

**Spoken or Written:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Spoken	10,409,858	<a href="#">344</a>	126/908	33.05
Written	87,903,571	<a href="#">2,269</a>	765/3,140	25.81
<b>total</b>	<b>98,313,429</b>	<b>2,613</b>	<b>891/4,048</b>	<b>26.58</b>

**Derived text type:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Fiction and verse	16,143,913	<a href="#">1,143</a>	312/452	70.8
Spoken conversation	4,233,962	<a href="#">263</a>	72/153	62.12
Other published written material	17,924,109	<a href="#">668</a>	212/710	37.27
Newspapers	9,412,174	<a href="#">153</a>	74/486	16.26
Other spoken material	6,175,896	<a href="#">81</a>	54/755	13.12
Non-academic prose and biography	24,178,674	<a href="#">232</a>	114/744	9.6
Unpublished written material	4,466,673	<a href="#">30</a>	22/251	6.72
Academic prose	15,778,028	<a href="#">43</a>	31/497	2.73
<b>total</b>	<b>98,313,429</b>	<b>2,613</b>	<b>891/4,048</b>	<b>26.58</b>

**Text type:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Demographically sampled	4,233,962	<a href="#">263</a>	72/153	62.12
Written books and periodicals	79,187,792	<a href="#">2,111</a>	709/2,684	26.66
Written miscellaneous	7,437,161	<a href="#">154</a>	53/421	20.71
Context-governed	6,175,896	<a href="#">81</a>	54/755	13.12
Written-to-be-spoken	1,278,618	<a href="#">4</a>	3/35	3.13
<b>total</b>	<b>98,313,429</b>	<b>2,613</b>	<b>891/4,048</b>	<b>26.58</b>

Your query "pretty\_AJO" returned 2613 hits in 891 different texts (98,313,429 words [4,048 texts]; frequency: **26.58 instances per million words**)

### - Standard query: pretty\_AV0 (Adverb)

Your query "pretty\_AV0" returned 4921 hits in 1508 different texts. The current solution set was found in 1508 texts. [View results](#)

Categories:

Categories (for crosstabs only):

**The following distribution was found:**

**Spoken or Written:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Spoken	10,409,858	<a href="#">1,113</a>	411/908	106.92
Written	87,903,571	<a href="#">3,808</a>	1,097/3,140	43.32
<b>total</b>	<b>98,313,429</b>	<b>4,921</b>	<b>1,508/4,048</b>	<b>50.05</b>

**Derived text type:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Other spoken material	6,175,896	<a href="#">691</a>	302/755	111.89
Spoken conversation	4,233,962	<a href="#">422</a>	109/153	99.67
Fiction and verse	16,143,913	<a href="#">1,511</a>	300/452	93.6
Other published written material	17,924,109	<a href="#">1,027</a>	278/710	57.3
Newspapers	9,412,174	<a href="#">376</a>	141/486	39.95
Unpublished written material	4,466,673	<a href="#">167</a>	59/251	37.39
Non-academic prose and biography	24,178,674	<a href="#">617</a>	252/744	25.52
Academic prose	15,778,028	<a href="#">110</a>	67/497	6.97
<b>total</b>	<b>98,313,429</b>	<b>4,921</b>	<b>1,508/4,048</b>	<b>50.05</b>

**Text type:**

Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Context-governed	6,175,896	<a href="#">691</a>	302/755	111.89
Demographically sampled	4,233,962	<a href="#">422</a>	109/153	99.67
Written books and periodicals	79,187,792	<a href="#">3,535</a>	1,000/2,684	44.64
Written miscellaneous	7,437,161	<a href="#">234</a>	77/421	31.46
Written-to-be-spoken	1,278,618	<a href="#">39</a>	20/35	30.5
<b>total</b>	<b>98,313,429</b>	<b>4,921</b>	<b>1,508/4,048</b>	<b>50.05</b>

Your query "pretty\_AV0" returned 4921 hits in 1508 different texts (98,313,429 words [4,048 texts]; frequency: **50.05 instances per million words**)

Result: More hits for adverbs than for adjectives found.