

# A8) "Pretty": adjective vs. adverb (BNC + BYU interface)

Martin Schneidmadel

Corpus: British National Corpus (BNC)

Programs: BYU-BNC

Tasks:

- *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.
- Carry out manual spot checks to find out to what extent the POS-tagging is adequate.

Notes:

- The corpus and the BYU interface can be accessed on the Internet: <http://corpus.byu.edu/bnc/>
- After a few searches you will have to create a username and password (free of charge).
- Users taking part in the course "Methods and Theories in Linguistics" can use the password **metheling** to join the account for the University of Bamberg: In order to do so, make sure the [Organization] in the form under "Personal Information" at the bottom of that page matches the name Universität Bamberg (exactly), and click [Update]. Then enter the password **metheling** and click on [Join Group], after which you will then be added to the group account for our university. You will then have increased access and advanced features.
- Try out the different display modes (upper left corner).
- If you want to save your results, the best idea seems to be highlighting and copying/pasting of content.
- The website itself includes a help function and ample documentation.

## Description of the Procedure:

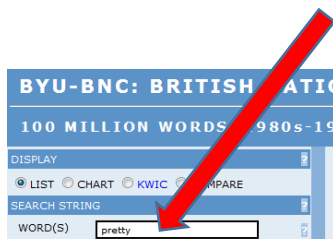
1. "Pretty" premodifying nouns.....	02
2. "Pretty" in predicative function.....	05
3. "Pretty" before Adjectives and other Adverbs.....	06
4. Spot Checks.....	07
5. Final Results.....	07

# 1."Pretty", premodifying nouns

Open the website: <http://corpus.byu.edu/bnc/>

The screenshot shows the BYU-BNC website interface. The top navigation bar includes the site name "BYU-BNC: BRITISH NATIONAL CORPUS" and the user name "M SCHNEIDMADEL". Below this, it states "100 MILLION WORDS, 1980s-1993" and a "-- START --" dropdown menu. The left sidebar contains a "DISPLAY" section with radio buttons for "LIST", "CHART", "KWIC", and "COMPARE". The "SEARCH STRING" field is empty. Below it are "COLLOCATES" and "POS LIST" sections. The "SECTIONS" section shows two "IGNORE" lists with various categories like "SPOKEN", "FICTION", "MAGAZINE", "NEWSPAPER", and "NON-ACAD". The "SORTING AND LIMITS" section has "SORTING" set to "FREQUENCY" and "MINIMUM" set to "FREQUENCY" with a value of "5". The main content area features a list of resources: "Full-text", "Wikipedia corpus (NEW)", "Word and Phrase (analyze texts)", "Word and Phrase (frequency lists)", "Word Frequency", and "Collocates". Each resource has a brief description of its capabilities.

Enter the word *pretty*.



This screenshot is identical to the previous one, but with the word "pretty" entered in the "WORD(S)" search field. A red arrow points to this field from the text above. The rest of the interface, including the sidebar and resource list, remains the same.

We have to find *pretty* in two different positions within the sentence:

- premodifying nouns/ in a predicative function: *pretty* as adjective
- before adjectives and other adverbs: *pretty* as adverb

The Website offers us different options for the position in the sentence in the Section **Pos List**. We are interested in *pretty* as premodifier of nouns.

BYU-NC: BRITISH NATIONAL CORPUS  
100 MILLION WORDS, 1980s-1990s

DISPLAY:  LIST  KWIC  COMPARE

SEARCH SETTINGS

WORD(S): pretty

COLLOCATIONS: [n\*] 4 4

POS LIST: noun

RANDOM SEARCH RESET

SECTIONS: SHOW

1 IGNORE  
-----  
SPOKEN  
FICTION  
MAGAZINE  
NEWSPAPER  
NON-ACAD

2  
-----  
SPO  
FICTI  
MAGAZI  
NEWSPAP  
NON-ACAD

SORTING AND LIMITS

SORTING: FREQUENCY

MINIMUM: FREQUENCY 5

CLICK TO SEE OPTIONS

Therefore we choose [n\*]

But we have to limit the number of words around *pretty*. Otherwise the corpus shows sentences like: “...the young girls were very pretty...” or “...the girls were pretty tired”.

# BYU-BNC: BRITISH NATIONAL CORPUS

100 MILLION WORDS, 1980s-1993

DISPLAY ?

LIST  CHART  KWIC  COMPARE

SEARCH STRING ?

WORD(S)  ?

COLLOCATES \*   ?

POS LIST ?

SECTIONS  SHOW

1

2

SORTING AND LIMITS ?

SORTING  ?

MINIMUM    ?

CLICK TO SEE OPTIONS ?

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT]

	<input type="checkbox"/>	CONTEXT
1	<input type="checkbox"/>	GIRL
2	<input type="checkbox"/>	FACE
3	<input type="checkbox"/>	WOMAN
4	<input type="checkbox"/>	POLLY
5	<input type="checkbox"/>	GIRLS
6	<input type="checkbox"/>	VILLAGE
7	<input type="checkbox"/>	SIGHT
8	<input type="checkbox"/>	VILLAGES
9	<input type="checkbox"/>	LADY
10	<input type="checkbox"/>	GARDEN
11	<input type="checkbox"/>	BOY

INTRODUCTION

[ COMPARE BNC AND COCA ] [ COMPARE SIDE-B

This website allows you to quickly and easily search the British National Corpus (BNC) in the 1980s - early 1990s, and now exists in

If you find this version of the BNC useful, you may also be interested in the 450 million word Corpus of Contemporary American

With the Collocation 0 / 1 ... we allow *pretty* only to stand in front of one noun instead of a construction where the noun stands before *pretty*.

The results:

BYU-BNC: BRITISH NATIONAL CORPUS

100 MILLION WORDS, 1980s-1993

M SCHNEIDMADEL !

history | lists | logout

DISPLAY ?

LIST  CHART  KWIC  COMPARE

SEARCH STRING ?

WORD(S)  ?

COLLOCATES \*   ?

POS LIST ?

SECTIONS  SHOW

1

2

SORTING AND LIMITS ?

SORTING  ?

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

COMPARE

	<input type="checkbox"/>	CONTEXT	FREQ	ALL	%	MI	
1	<input type="checkbox"/>	GIRL	87	14383	0.60	8.72	
2	<input type="checkbox"/>	FACE	68	32434	0.21	7.19	
3	<input type="checkbox"/>	WOMAN	46	21791	0.21	7.20	
4	<input type="checkbox"/>	POLLY	39	786	4.96	11.75	
5	<input type="checkbox"/>	GIRLS	35	8950	0.39	8.09	
6	<input type="checkbox"/>	VILLAGE	34	10934	0.31	7.76	
7	<input type="checkbox"/>	SIGHT	27	6281	0.43	8.23	
8	<input type="checkbox"/>	VILLAGES	19	2735	0.69	8.92	
9	<input type="checkbox"/>	LADY	19	9215	0.21	7.17	
10	<input type="checkbox"/>	GARDEN	18	10726	0.17	6.87	
11	<input type="checkbox"/>	BOY	17	12560	0.14	6.56	

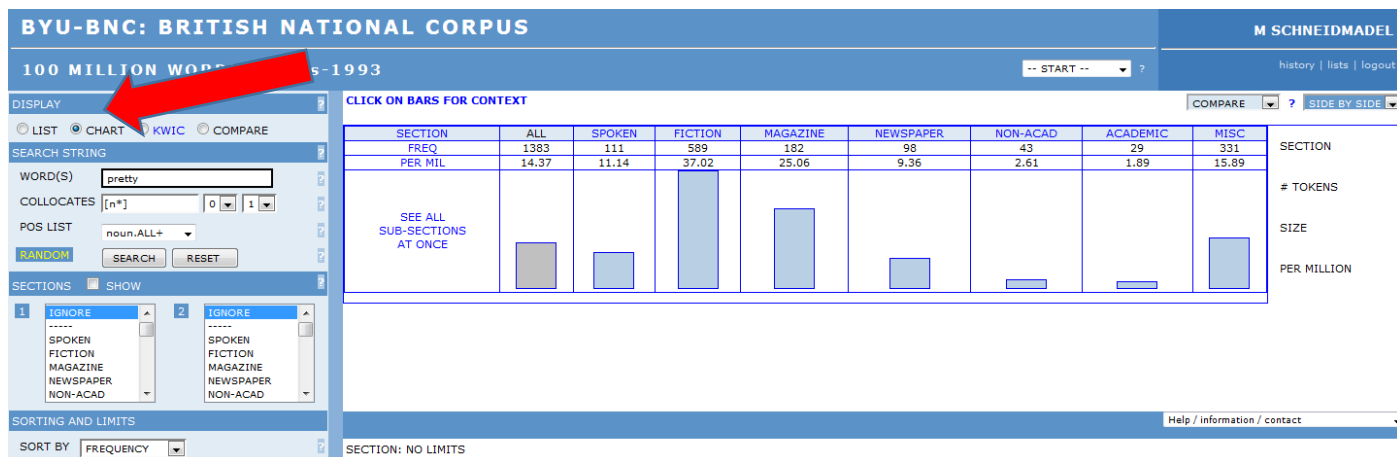
OVERVIEW: MARTIN SCHNEIDMADEL

Help / information / contact

Right now we know the nouns which are highly frequently premodified by *pretty*: girl, face, woman, things...

But the task is to find the frequencies of these grammatical functions **in different sections**, not before certain words!

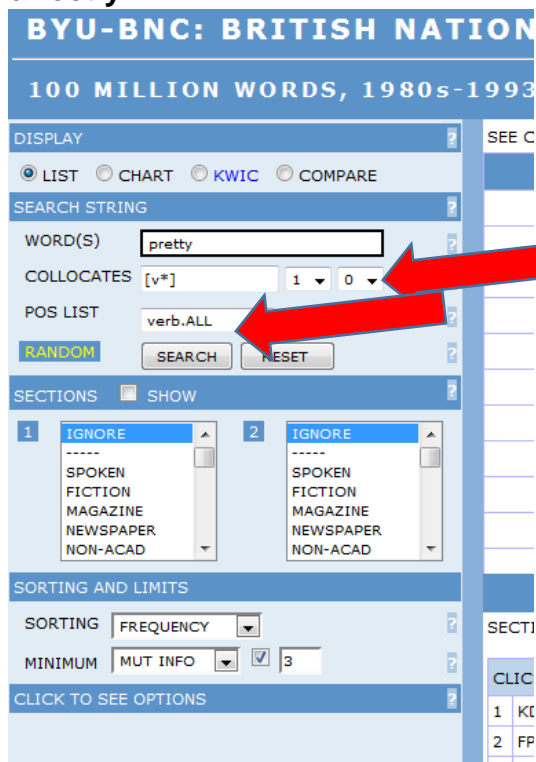
For this purpose we choose **CHART** instead of **LIST** and we can see in which sections *pretty* occurs as a premodifier of nouns.



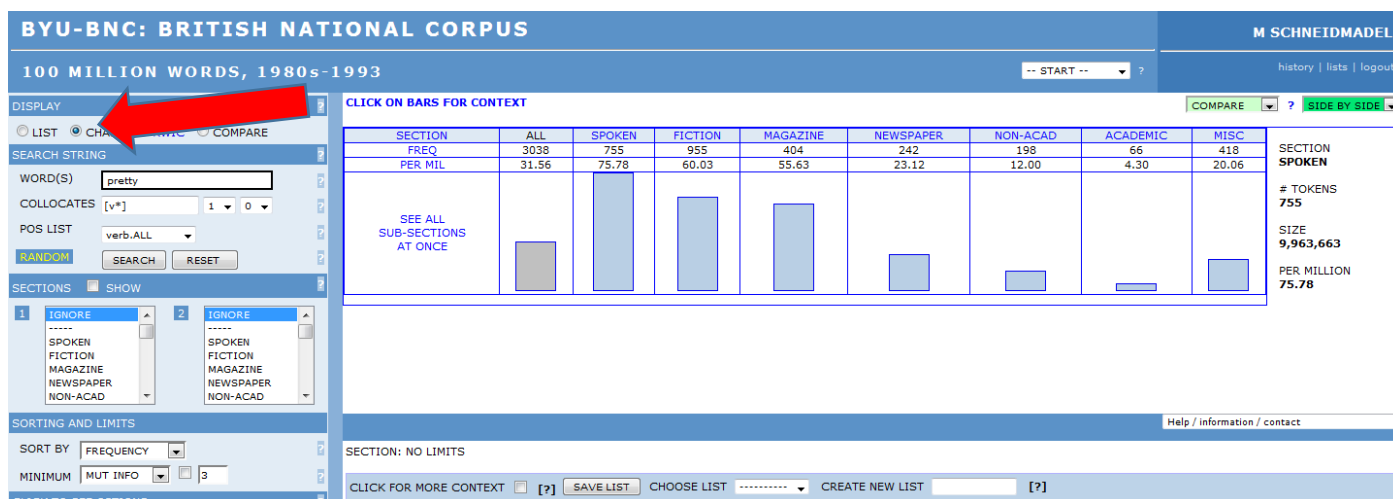
## 2. “Pretty” in predicative function

However, it’s not finished yet! What about *pretty* in predicative position? Just change from [n\*] to verb [v\*] / verb. ALL in **POS LIST**.

Take care the parameter is now 1 / 0 in **COLLOCATES** so that *pretty* follows the verb directly.



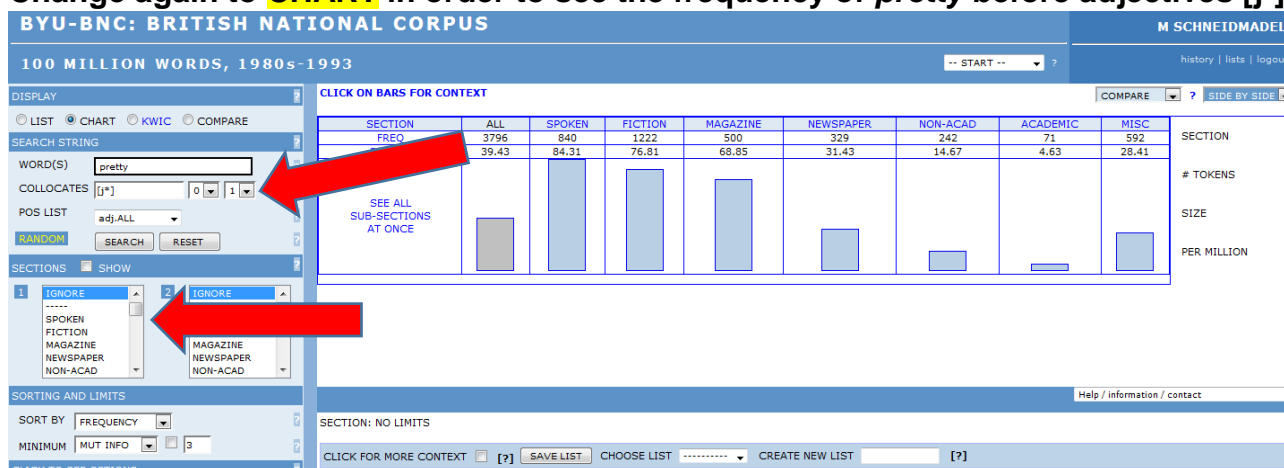
Change to **CHART** and see the results.



### 3. "Pretty" before Adjectives and other Adverbs

Right now, we do the same procedure for pretty (adv) before adjectives and other adverbs. We choose [j\*] for adjectives and [r\*] for adverbs in the **POS LIST**. Again we have to check the range at **COLLOCATES**, in this case 0 / 1.

Change again to **CHART** in order to see the frequency of *pretty* before adjectives [j\*]:



and before adverbs [r\*]:

#### 4. Spot Checks

We have to carry out manual spot checks to find out to what extent the POS-tagging is adequate. For this purpose, go back to **LIST**. Then click on the words in the list and read the sentences.

My spot checks have shown that the POS-tagging works quite well with “pretty” as adjective and adverb, maybe because we are not looking for a very complex structure.

However, in the SPOKEN-sources, there are some unclear sentence patterns which make it almost impossible to understand the sentence and understand the use of *pretty* in the sentence. Here some examples:

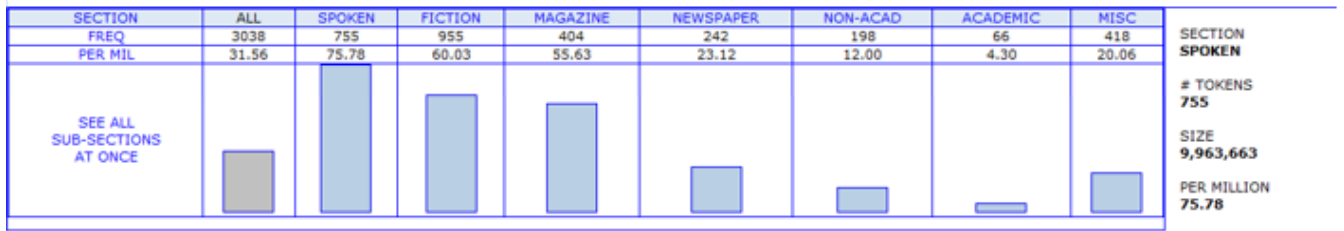
3	J3M	S_meeting	A B C	it's a precedent of our (unclear) (SP:J3MPSUNK) (unclear) (SP:PS3M7) Erm (pause) it's a <b>pretty close</b> run thing but by a narrow majority,
4	J3M	S_meeting	A B C	Er we'd like to thank Bert, too, for (pause) standing. (unclear) <b>pretty close</b> run thing I can assure you. (SP:PS3MA) (unclear) historical int
5	JAD	S_pub_debate	A B C	Now I think that er the County Council have reworked their figures and are probably <b>pretty close</b> to us now on on these matters. Erm per
6	KJT	S_sportslive	A B C	both Ian Botham who's at second (unclear) and Graham Gooch, they both stand <b>pretty close</b> , they work on the theory that it's better to di

#### 5. Final Results:

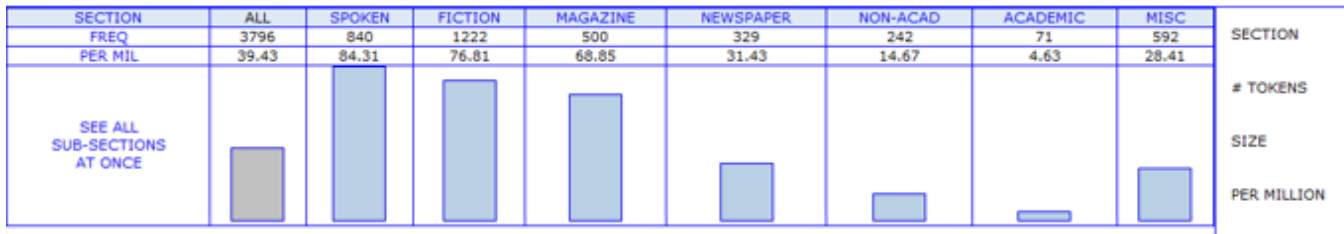
part of a noun phrase:

SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACAD	ACADEMIC	MISC	
FREQ	1383	111	589	182	98	43	29	331	SECTION
PER MIL	14.37	11.14	37.02	25.06	9.36	2.61	1.89	15.89	# TOKENS
									SIZE
									PER MILLION

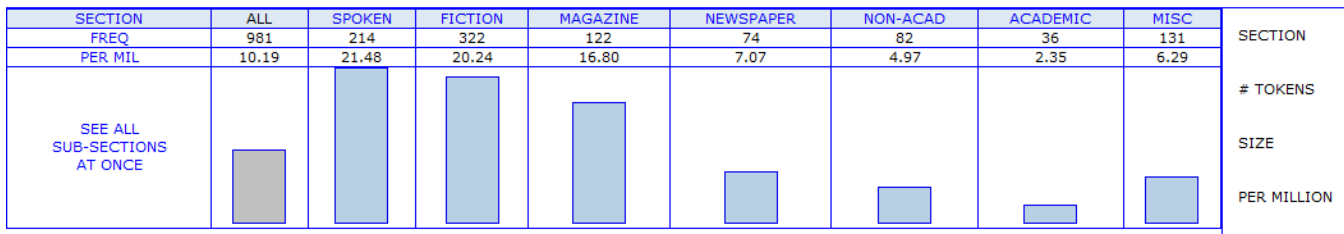
**in predicative function:**



**before adjectives:**



**and adverbs:**



**Source: British National Corpus (BNC)**