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

Corpus: British National Corpus (BNC)

Programs: BYU-BNC

<u>Tasks</u>:

- *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: <u>http://corpus.byu.edu/bnc/</u>
- 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	
3. "Pretty" before Adjectives and other Adverbs	
4. Spot Checks	07
5. Final Results	07

1."Pretty", premodifying nouns

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

BYU-BNC: BRITISH NAT	IONAL CORPU	S	M SCHNEIDM/	ADE				
00 MILLION WORDS, 1980s-	1993		START 💌 🤊 history lists	logo				
SPLAY								
LIST O CHART O KWIC O COMPARE			urces that are based on the BYU corpora:					
ARCH STRING	Full-text		million words of full-text data for COCA (190,000 texts), or 1.8 billion words for GloWbE (1,800,000 texts). With this data, you texts from the corpora on your own computer, rather than having to use the web interface.					
ORD(S)	Wikipedia corpus	Quickly and eas	sily create "virtual" corpora from the 4.4 million articles of Wikipedia (1.9 billion words) on almost any topic biology,					
DILIOCATES 2 DIS LIST 2	(NEW)	investments, ca corpora.	rs, Buddhism, etc. Search these virtual corpora, compare them to each other, and create keyword/frequency lists from your					
TIONS SHOW	Word and Phrase (analyze texts)							
IGNORE								
SPOKEN SPOKEN FICTION FICTION MAGAZINE MAGAZINE NEWSPAPER NEWSPAPER	Word Frequency	Word Frequency You can also download lists showing the frequency of the top 60,000 lemmas by genre (and sub-genre). Free list of COCA. Download the 100,000 integrated word list from COCA, COHA, BNC, and SOAP the largest, corrected freq						
NON-ACAD V NON-ACAD V	Collocates	Download lists v	with the top 200-300 collocates (nearby words) for 60,000 different lemmas 4,300,000 node/collocate pairs in all.					
TING AND LIMITS	OVERVIEW: MARTIN	SCHNEIDMADEL	Help / information / contact					
RTING FREQUENCY								
IIMUM FREQUENCY 💌 🗖 5	Personal information	on						
K TO SEE OPTIONS	Institution		University of Bamberg Update					
	Country		GERMANY					
	Support							
	Status		It looks like you haven't yet donated to help support the BYU corpora. Please consider doing so now.					
	Other recent users fro	om your university	UNIVERSITAET BAMBERG Schlitter, Müller, Knappe, Lucas, Schützler, Werner, Christel, Denzler, Vetter, Vetter, Bickert, Weiler, hanif, Jokiel, Buttafoco, Mohr, Leisgang, Buckreus, Gläsel, Deininger, Flieger, Schneider, Schreiber, Viktor, Mohr, Fleischmann, Sheppard, Dinkel, Marschner, JANG, Roth, Feng, Schierwagen, Nguyen, Kroll, Schneidmadel					
			It looks like there are a number of other recent users from your university or college. Please consider the possibility of an institution-wide license to help support the BYU corpora. This will provide all users from your university with additional access and features. (And it's just the right thing to do if you're using the corpora that much.)					
	Activity							
	Last login (before this	time)	10/29/2015 10:16:09 AM (Utah time; same as Denver)					

Enter the word *pretty*.

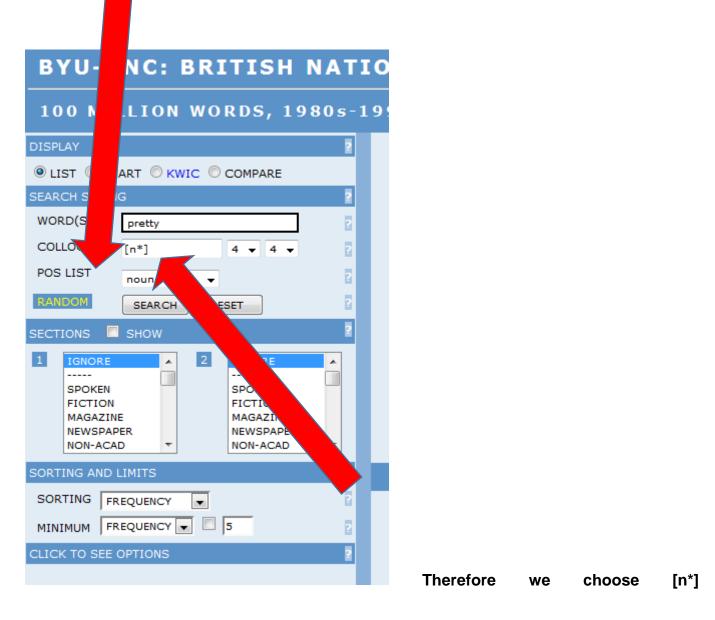
7

YU-BNC: BRITISH ATI	IONAL CORPUS	5				M SCHNEIDM		
.00 MILLION WORDS 980s-1	.993				START 🗸	? history lists		
LIST CHART KWIC			rces that are based on the BYU o					
ARCH STRING	Full-text				i), or 1.8 billion words for GloWbE (1,8 rather than having to use the web inter			
ORD(S) pretty	Wikipedia corpus	Quickly and easi	ly create "virtual" corpora fro	m the 4.4 million art	icles of Wikipedia (1.9 billion words) o	n almost any topic biology.		
DLLOCATES	(NEW)	(NEW) investments, cars, Buddhism, etc. Search these virtual corpora, compare them to each other, and create keyword/frequ						
S LIST 2		corpora.						
NDOM SEARCH RESET ?	Word and Phrase (analyze texts)							
	Word and Phrase (frequency lists)	page) definition, frequency by I links from one resource to another.						
SPOKEN SPOKEN FICTION FICTION MAGAZINE MAGAZINE NEWSPAPER NEWSPAPER	Word Frequency				lemmas by genre (and sub-genre). Fre , BNC, and SOAP the largest, correc			
NON-ACAD T	Collocates	Download lists w	ith the top 200-300 collocates (n	earby words) for 60	.000 different lemmas 4,300,000 no	de/collocate pairs in all.		
TING AND LIMITS	OVERVIEW: MARTIN S	CHNEIDMADEL				Help / information / contact		
RTING FREQUENCY								
IMUM FREQUENCY S	Personal information	_						
K TO SEE OPTIONS	Institution	n	University of Bamberg	Update				
					D OTHER USERS FROM GERMANY			
	Country		GERMANY -	Update	D OTHER USERS FROM GERMANT			
	Support Status		It looks like you haven't vet do	nated to bein supp	ort the BYU corpora. Please conside	r doing so now		
	Other recent users from	m vour university	UNIVERSITAET BAMBERG	need to need suppr	in the pro corpora ricebe conside			
		,,	Schlützer, Müller, Knappe, Lucas, Schützler, Werner, Christel, Denzler, Vetter, Vetter, Bickert, Weiler, hanif, Jokiel, Buttafoco, Nohr, Leisgang, Buckreus, Gläsel, Deininger, Fleger, Schneider, Schreiber, Viktor, Nohr, Fleischmann, Sheppard, Dinkel, Marschner, JANS, Roth, Feng, Schierwagen, Nguyen, Kroll, Schneidmadel					
			**** 1 PT -11		ra from your university or college. Pla	ase consider the possibility of an		
			institution-wide license to h	elp <u>support the BY</u>	U corpora. This will provide all users do if you're using the corpora that mu	from your university with additional		
	Activity		institution-wide license to h	elp <u>support the BY</u>	U corpora. This will provide all users	from your university with additional		

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.



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 NAT	IONAL	CORP	US
100 MILLION WORDS, 1980s-	1993		
DISPLAY	SEE CONTEXT	CLICK ON W	VORD OR SELECT WORDS + [CC
● LIST ○ CHART ○ KWIC ○ COMPARE			CONTEXT
SEARCH STRING	1		GIRL
WORD(S) pretty [n*]	2		FACE
COLLOCATES * 0 • 1 • 2	3		WOMAN
POS LIST	4		POLLY
RANDOM SEARCH RESET	5		GIRLS
SECTIONS SHOW	6		VILLAGE
1 IGNORE 2 IGNORE	7		SIGHT
	8		VILLAGES
FICTION FICTION MAGAZINE MAGAZINE	9		LADY
NEWSPAPER NEWSPAPER	10		GARDEN
NON-ACAD T	11		воу
SORTING AND LIMITS	INTRO	DUCTION	
SORTING FREQUENCY			
MINIMUM MUT INFO 💌 🗹 3	[CON	IPARE BNC A	ND COCA] [COMPARE SIDE-E
CLICK TO SEE OPTIONS			you to quickly and easily sear - early 1990s, and now exists in
			sion of the BNC useful, you may

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*.

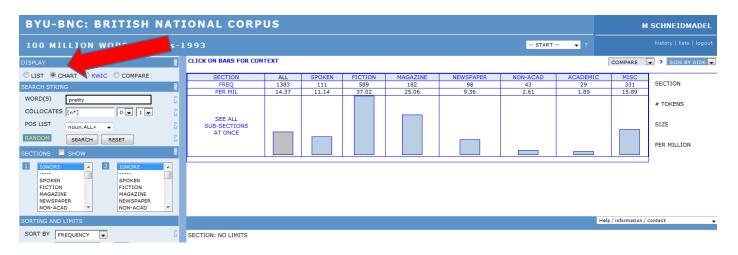
The results:

100 MILLION WORDS, 1980s-1993											
SEE CONTEXT	CLICK ON	WORD OR SELECT WORDS + [CONTEXT] [HELP]				COMPARE 💌 ?	SIDE BY SIDE				
		CONTEXT	FREQ								
1		GIRL	87	14383	0.60	8.72					
2		FACE	68	32434	0.21	7.19					
3		WOMAN	46	21791	0.21	7.20					
4		POLLY	39	786	4.96	11.75					
5		GIRLS	35	8950	0.39	8.09	_				
6		VILLAGE	34	10934	0.31	7.76					
7		SIGHT	27	6281	0.43	8.23	-				
8		VILLAGES	19	2735	0.69	8.92	-				
9		LADY	19	9215	0.21	7.17	-				
10		GARDEN	18	10726	0.17	6.87	-				
11		BOY	17	12560	0.14	6.56					
	SEE CONTEXT	SEE CONTEXT: CLICK ON 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 -	SEE CONTEXT: LICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP] CONTEXT CONTEXT 1 GIRL GIRL 2 FACE GIRL 3 WOMAN GIRLS 6 VILLAGE GIRL 7 SIGHT SIGHT 8 VILLAGES GIRLS 9 LADY GARDEN 11 BOY GIRLS	SEE CONTEXT: LICK ON VERLO OR SELECT WORDS + [CONTEXT] [HELP] I CONTEXT FREQ 1 CIRL 87 2 FACE 68 3 WOMAN 46 4 POLLY 39 5 GIRLS 35 6 VILLAGE 34 7 SIGHT 27 8 VILLAGES 19 10 GARDEN 18 11 BOY 17	SEE CONTEXT: CONTEXT FREQ ALL 1 CONTEXT FREQ ALL 1 CONTEXT FREQ ALL 1 CONTEXT FREQ ALL 1 CIRL FREQ ALL 2 FACE FREQ 532434 3 WOMAN 466 2 4 POLLY 39 2 5 GIRLS 355 8950 6 VILLAGE 344 10934 7 SIGHT 27 2 8 VILLAGES 19 2 2735 9 LADY 19 9 9215 10 GARDEN 18 10226	SEE CONTEXT: CONTEXT FREQ ALL % 1 CONTEXT FREQ ALL % 2 FACE 668 32434 0.21 3 WOMAN 466 21791 0.21 4 POLLY 39 4 4.96 5 GIRLS 35 950 0.39 6 VILLAGE 34 10934 0.31 7 SIGHT 27 2 6211 0.43 8 VILLAGES 19 2 2735 0.69 9 LADY 19 9215 0.21 10 GARDEN 18 9215 0.14	CONTEXT (HELP) COMPARE or ? Image: Context or Select words + [CONTEXT] (HELP) FREQ ALL % MI 1 GIRL 67 14383 0.60 3.72 2 FACE 68 32434 0.21 7.93 3 WOMAN 46 21791 0.21 7.20 4 POLY 39 2 786 4.96 11.75 5 GIRLS 35 3 8950 0.39 8.09 6 VILLAGE 34 2 10934 0.31 7.76 7 SIGHT 27 2 6281 0.43 8.23 8 VILLAGES 19 2 75 0.69 8.92 9 LADY 19 2 9215 0.21 7.17 10 GARDEN 18 10726 0.17 6.87				

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!

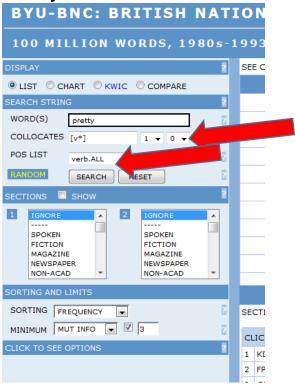
For this purpose we choose CHART instead of LIST and we can see in which sections pretty occurs as a premodifier of nouns.



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.

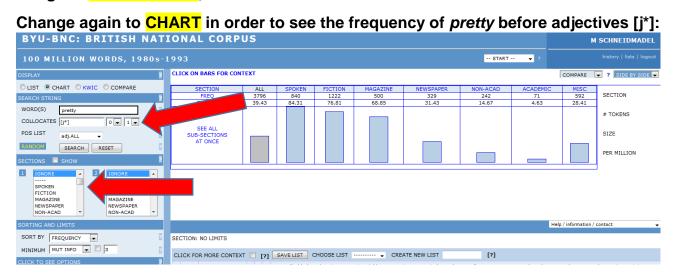


Change to CHART and see the results.

LIST Other Incompare EARCH STRING Pictor ALL SPOKEN FICTION MAGAZINE NON-ACAD ACADEMIC Misc WORD(S) pretty Collocates (v*) 1 0 Pictor ALL SPOKEN FICTION MAGAZINE NON-ACAD ACADEMIC Misc SECTION Section <th< th=""><th>BYU-BNC: BRITISH NAT</th><th></th><th colspan="2">M SCHNEIDMADEL</th></th<>	BYU-BNC: BRITISH NAT		M SCHNEIDMADEL										
Cultor Compare	100 MILLION WORDS, 1980s-	100 MILLION WORDS, 1980s-1993											
EARCH STRING FREQ 3038 755 955 404 108 66 418 SECTION WORD(S) pretty collocates [v*] 1 0 75.78 60.03 55.63 23.12 12.00 4.30 20.06 spoken POS LIST verb.all verb.all SEARCH RESET SEE SUB-SECTIONS State 55.63 23.12 12.00 4.30 20.06 spoken FOOLOCATES [v*] 1 0 Tonce 0 0 55.63 23.12 12.00 4.30 20.06 spoken FORMUSED SEARCH RESET Image: Collocates Sections State	DISPLAY	CLICK ON BARS FOR CON	TEXT							COMPARE	• ? SIDE BY SIDE •		
EARCH SIRING PRE MIL 0.06 75.73 60.03 55.63 23.12 12.00 4.30 20.06 SPOKEN VORD(S) pretty I 0 75.73 60.03 55.63 23.12 12.00 4.30 20.06 SPOKEN POS LIST verb.ALL SEARCH RESET Image: Search SEE Image: Search SEE	CIIST CHA	SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACAD	ACADEMIC	MISC			
WORD(S) pretty Image: Constrained and the provided and the provi	SEARCH STRING												
COLLOCATES [v*]		PER MIL	31.56	75.78	60.03	55.63	23.12	12.00	4.30	20.06	SPOKEN		
SEARCH RESET COTIONS SEARCH RESET COTIONS SEARCH RESET COTIONS SEARCH RESET COTIONS SHOW SECTIONS AT ONCE SPOKEN PER MILLION 75.78 SPOKEN PER MILLION 75.78 SPOKEN PER MILLION 75.78 SPOKEN PER MILLION 75.78 SPOKEN SPOKEN MAGAZINE NON-ACAD SPOKEN NON-ACAD SPOKEN NON-ACAD SPOKEN NON-ACAD SPOKEN NON-ACAD SPOKEN NON-ACAD SPOKEN	WORD(S) pretty 2										# TOKENS		
POS LIST verb.ALL SUB_SECTIONS EXTRODUCT SUB_SECTIONS ECTIONS SHOW I IONORE PICTION SORTEN NON-ACAD NON-ACAD ORTING AND LIMITS SECTION: NO LIMITS SORT BY REQUENCY SORT BY REQUENCY I SORT BY REQUENCY SECTION: NO LIMITS ECTION: NO LIMITS ECTION: NO LIMITS	COLLOCATES [v*] 1 V 0 V										755		
Verbault 9,963,663 PER MILLION SEARCH ECTIONS SHOW Image: Show	200 J 107										0175		
REALIZION SEARCH RESET PER MILLION ECTIONS SHOW PICTION PICTION I DENORE PICTION SPOKEN PICTION FICTION SPOKEN PICTION MAGAZINE NON-ACAD PICTION NON-ACAD NON-ACAD PICTION SORT BY FREQUENCY SECTION: NO LIMITS	verb.ALL 👻												
ECTIONS SHOW PER MILLION I IONORE SPOKEN SPOKEN SPOKEN FICTION MAGAZINE SPOKEN NON-ACAD NON-ACAD OKTING AND LIMITS SECTION: NO LIMITS SORT BY FREQUENCY SORT BY FREQUENCY 3 CLICK FOR MORE CONTEXT (PER MILLION 75.78 76.72 75.78 76.72 76.72 76.72 76.72 76.72 77.72 76.72 77.72 76.72 77.73 76.72 77.72 76.72 77.73 77.73 77.73 77.73 77.73 77.73 77.73 77.74 <p< td=""><td>RANDOM SEARCH RESET ?</td><td>ATONCE</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></p<>	RANDOM SEARCH RESET ?	ATONCE											
I GNORE I GNORE SPOKEN SPOKEN FICTION MAGAZINE NON-ACAD NON-ACAD ORTING AND LIMITS SORT BY FREQUENCY SORT BY FREQUENCY I GNORE SECTION: NO LIMITS SORT BY FREQUENCY I GNORE SECTION: NO LIMITS SORT BY FREQUENCY I GNORE SECTION: NO LIMITS SOLUCK FOR MORE CONTEXT [2] SAVE LIST CHOOSE LIST CREATE NEW LIST [2]													
SPOKEN	SECTIONS SHOW										75.78		
SORT BY FREQUENCY SECTION: NO LIMITS SECTION: NO LIMITS CLICK FOR MORE CONTEXT [] [] SAVE LIST CHOOSE LIST	SPOKEN FICTION MAGAZINE NEWSPAPER NEWSPAPER										-		
	SORTING AND LIMITS								н	elp / information /	contact 🗸		
CLICK FOR MORE CONTEXT [] SAVE LIST CHOOSE LIST CHOOSE LIST CHOOSE LIST []	SORT BY FREQUENCY	SECTION: NO LIMITS											
CLICK FOR MORE CONTEXT [] SAVE LIST CHOOSE LIST CHOOSE LIST CHOOSE LIST []	MINIMUM MUT INFO												
		CLICK FOR MORE CONTEX	T 🔳 [?] 🗌	SAVE LIST	CHOOSE LIST	CRE	ATE NEW LIST	[?]					

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.



and before adverbs [r*]:

BYU-BNC: BRITISH NA	TIONAL CORP	US							N	I SCHNEIDMADEI
100 MILLION WORDS, 1980s	5-1993						START	• ?		history lists logou
ISPLAY	2 CLICK ON BARS FOR CON	EXT		CLI	CK ON COLUMN H	EADINGS FOR FREQ	UENCY IN SUB-SEC	CTION	COMPARE	SIDE BY SIDE
LIST CHART KWIC COMPARE		ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACAD	ACADEMIC	MISC	
EARCH STRING		981	214	322	122	74	82	36	131	SECTION
/ORD(S) pretty [r*]	Sec MIL	10.19	21.48	20.24	16.80	7.07	4.97	2.35	6.29	NEWSPAPER
										# TOKENS
OLLOCATES 0 - 1 -	SEE ALL									74
DS LIST adv.ALL -	2 SUB-SECTIONS									SIZE
ANDOM SEARCH RESET	AT ONCE									10,466,422
										PER MILLION
CTIONS 🔲 SHOW	8									7.07
NEWCOADED NEWCOADED										
NEWSPAPER NON-ACAD T NON-ACAD T	2							н	els / information /	contact
NON-ACAD T	2							Н	ielp / information /	contact
NON-ACAD NON-ACAD	E SECTION: NO LIMITS							н	lelp / information /	contact
NON-ACAD NON-ACAD KTING AND LIMITS RTING AND LIMITS RT BY FREQUENCY MUT INFO	-	T	[?] SAVE	ELIST CHOOS	5E LIST		ST	[?]	lelp / information /	contact
NON-ACAD NON-ACAD	SECTION: NO LIMITS					✓ CREATE NEW LIS ay oh it's it's pretty c		[3]		
NON-ACAD NON-ACAD	SECTION: NO LIMITS	A	A B C Mr () (unclear) if i	's spelt you may sa		lose. It's wrong you	[?] see. Another one	e. You might say	/ ham
NON-ACAD NON-ACAD KTING AND LIMITS RT BY FREQUENCY MUT INFO 3	CLICK FOR MORE CONTEX	A A	A B C Mr (A B C can't w) (unclear) if it vork it out. I rea	's spelt you may sa lly can't afford (SP:	ay oh it's it's <u>pretty</u> c	lose. It's wrong you etty close. So you	[?] I see. Another one don't need all this	a. You might say	/ ham do you
NON-ACAD NON-ACAD	CLICK FOR MORE CONTEX 1 KPA S_conv 2 KSV S_conv	A A A	ABCMr(ABCcan'tw ABCit'sap) (unclear) if it vork it out. I rea recedent of our	's spelt you may sa lly can't afford (SP: (unclear) (SP:J3MP	ay oh it's it's <u>pretty c</u> PS1BY) It must be <u>pr</u>	lose. It's wrong you etty close. So you PS3M7) Erm (pause)	[?] u see. Another one don't need all this it's a <u>pretty clo</u>	a. You might say s, surely? What <u>se</u> run thing bu	r ham do you t by a narrow majorit
NON-ACAD NON-ACAD KTING AND LIMITS RT BY FREQUENCY MUT INFO 3	CLICK FOR MORE CONTEX 1 KPA S_conv 2 KSV S_conv 3 J3M S_meeting	م م م م	A B C Mr (A B C can't w A B C it's a p A B C Er we'd) (unclear) if it vork it out. I rea recedent of our d like to thank B	's spelt you may sa lly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause	ay oh it's it's <u>pretty c</u> PS1BY) It must be <u>pr</u> SUNK) (unclear) (SP:	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi	[?] i see. Another one don't need all this it's a <u>pretty clo</u> ing I can assure y	e. You might say s, surely? What <u>se</u> run thing bu rou. (SP:PS3MA	/ ham do you t by a narrow majorit) (unclear) historical i
NON-ACAD NON-ACAD	CLICK FOR MORE CONTEX KPA S_conv KSV S_conv 3 J3M S_meeting 4 J3M S_meeting	۹ ۹ ۹ ۹	A B C Mr (A B C can't w A B C it's a p A B C Er we'c A B C Now I) (unclear) if it vork it out. I rea recedent of our d like to thank B think that er the	's spelt you may sa lly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause County Council ha	ay oh it's it's <u>pretty c</u> PS1BY) It must be <u>pr</u> SUNK) (unclear) (SP:) standing. (unclear) j	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi ures and are probabl	[?] i see. Another one don't need all this it's a <u>pretty clo</u> ing I can assure y ly <u>pretty close</u> to	a. You might say s, surely? What <u>se</u> run thing bu rou. (SP:PS3MA o us now on on	/ ham do you t by a narrow majoril) (unclear) historical these matters. Erm p
NON-ACAD NON-ACAD TTING AND LIMITS RT BY FREQUENCY MIMUM MUT INFO 3	SECTION: NO LIMITS CLICK FOR MORE CONTEX 1 KPA S_conv 2 KSV S_conv 3 J3M S_meeting 4 J3M S_meeting 5 JAD S_pub_debate	۹ ۹ ۹ ۹ ۹	A B C Mr (A B C can't w A B C it's a p A B C it's a p A B C Er we'c A B C Now I A B C both Ia) (unclear) if it vork it out. I rea recedent of our d like to thank B think that er the an Botham who's	's spelt you may se lly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause County Council ha s at second (unclear	ay oh it's it's <u>pretty c</u> PS1BY) It must be <u>pr</u> (SUNK) (unclear) (SP:) standing. (unclear) j ve reworked their figu	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi ures and are probabl , they both stand <u>pr</u>	[?] u see. Another one don't need all this) it's a <u>pretty clo</u> ing I can assure y ly <u>pretty close</u> to <u>retty close</u> , they	a. You might say s, surely? What <u>se</u> run thing bu ou. (SP:PS3MA) o us now on on work on the the	y ham do you t by a narrow majoril ((unclear) historical these matters. Erm p cory that it's better to
NON-ACAD NON-ACAD	SECTION: NO LIMITS CLICK FOR MORE CONTEX 1 KFA S.conv 2 KSV S.conv 3 J3M S.meeting 4 J3M S.meeting 5 JAD S.pub_debate 6 KJT S.sportslive	۹ ۹ ۹ ۹ ۹ ۹ ۹	A B C Mr (A B C can't w A B C it's a p A B C new'c A B C both Ia A B C , but I') (unclear) if it vork it out. I rea recedent of our d like to thank B think that er the an Botham who's m so pleased he	spelt you may se ly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause County Council ha s at second (unclea s's back in Oxfords)	ay oh it's it's <u>pretty c</u> PS1BY) It must be <u>pr</u> (SUNK) (unclear) (SP:) standing. (unclear) j ve reworked their figu r) and Graham Gooch	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi ures and are probabl , they both stand <u>pr</u> ame <u>pretty close</u> a	[?] Jose, Another one don't need all this Jit's a <u>pretty clo</u> ing I can assure y ly <u>pretty close</u> to <u>etty close</u> , they it one time. This y	a. You might say is, surely? What se run thing bu rou. (SP:PS3MA) to us now on on work on the the rear, er been a	y ham do you t by a narrow majorii) (unclear) historical these matters. Erm p cory that it's better to bit iffy, and
NON-ACAD NON-ACAD	SECTION: NO LIMITS CLICK FOR MORE CONTEX KSV 5_conv KSV 5_conv KSV 5_conv J J3M 5_meeting JAD 5_pub_debate KAT 5_brotast_news KRT 5_brotast_news	A A A A A A Sn A	A B C Mr (A B C can't w A B C it's a p A B C it's a p A B C reve'c A B C now 1 ft A B C both 1 at A B C , but 1' A B C what tu) (unclear) if it vork it out. I rea recedent of our d like to thank B think that er the an Botham who's m so pleased he urns me on yet.	's spelt you may se lly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause County Council ha s at second (unclear s's back in Oxfordst But I'd say (SP:PS:	ay oh it's it's <u>pretty c</u> PSIBY) It must be <u>pr</u> SUNK) (unclear) (SP:) standing. (unclear) j ve reworked their figu r) and Graham Gooch nire. (SP:PS63K) He c	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi ures and are probabl , they both stand <u>pr</u> ame <u>pretty close</u> a 30) that's <u>pretty clos</u>	[?] u see. Another one don't need all this it's a <u>pretty clos</u> ing I can assure y <u>y pretty close</u> to <u>etty close</u> , they at one time. This y <u>use</u> . (break-in-reco	a. You might say a, surely? What <u>se</u> run thing bu rou. (SP:PS3MA) o us now on on work on the the rear, or been a ording) (advert)	/ ham do you by a narrow majorit) (unclear) historical these matters. Erm p sory that it's better to bit iffy, and) (SP:FXTPSUNK) (un
NON-ACAD NON-ACAD RTING AND LIMITS DRT BY FREQUENCY	SECTION: NO LIMITS CLICK FOR MORE CONTEX 1 KFA 5 Sconv 3 JM 5 JAD 6 KT 7 KRT 7 KRT 8 PAT 8 PAT	A A A A A A Sn A Sn A	A B C Mr (A B C can't w A B C it's ap A B C both lat A B C what tt A B C is not n') (unclear) if if work it out. I rea recedent of our d like to thank B think that er the an Botham who's m so pleased he urns me on yet. ru rude. I've h	's spelt you may se lly can't afford (SP: (unclear) (SP:J3MP ert, too, for (pause County Council ha s at second (unclean s's back in Oxfordsh But I'd say (SP:PS) eard people (SP:PS)	ay oh it's it's <u>pretty c</u> PS1BY) It must be pr SUNK) (unclear) (SP:) standing. (unclear) J ve reworked their figur r) and Graham Gooch nire. (SP:PS63K) He c 23R) (laugh) (SP:PS2)	lose. It's wrong you etty close. So you PS3M7) Erm (pause) pretty close run thi ures and are probabl , they both stand <u>pr</u> ame <u>pretty close</u> a 30) that's <u>pretty clo</u> S3DB) on er (SP:PS	[?] don't need all this don't need all this it's a <u>pretity clo</u> se ing I can assure y by <u>pretity close</u> , they it one time. This y <u>sse</u> . (break-in-rec 3DC) <u>Pretity Clos</u>	e. You might say ; surely? What <u>se</u> run thing bu' rou. (SP:PS3MA) o us now on on bus now on the the rear, er been a ording) (advert <u>se</u> . (SP:PS3DB)	y ham do you by a narrow majorit o) (unclear) historical i these matters. Erm p oroy that i's better to bit iffy, and (SP:FXTPSUNK) (un Yeah. (SP:PS3DC) N

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.

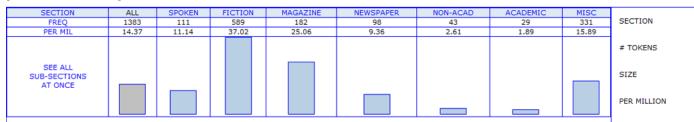
My spot checks have shown that the POS-tagging works quite well with "pretty" as adjective and adverb, maybe because we are not looking for a very complex structure.

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

5. Final Results:

part of a noun phrase:



in predicative function:

FREQ PER MIL	ALL 3038 31.56	55 755 75.78	FICTION 955 60.03	MAGAZINE 404 55.63	NEWSPAPER 242 23.12	NON-ACAD 198 12.00	ACADEMIC 66 4.30	MISC 418 20.06	SECTION SPOKEN
SEE ALL SUB-SECTIONS AT ONCE									# TOKENS 755 SIZE 9,963,663 PER MILLION 75,78

before adjectives:

SECTION FREQ PER MIL	ALL 3796 39.43	840 84.31	FICTION 1222 76.81	MAGAZINE 500 68.85	NEWSPAPER 329 31.43	NON-ACAD 242 14.67	ACADEMIC 71 4.63	MISC 592 28.41	SECTION
SEE ALL SUB-SECTIONS AT ONCE									# TOKENS SIZE PER MILLION

and adverbs:

	SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACAD	ACADEMIC	MISC	
	FREQ	981	214	322	122	74	82	36	131	SECTION
	PER MIL	10.19	21.48	20.24	16.80	7.07	4.97	2.35	6.29	
	SEE ALL SUB-SECTIONS AT ONCE									# TOKENS SIZE PER MILLION
_										

Source: British National Corpus (BNC)