

Outline

- What is specialist lexis?
- Making decisions about what lexis to teach
 - Corpus analysis tools
- Creating learning activities

What we mean by specialist lexis

Lexis:

- individual words
- combinations of words

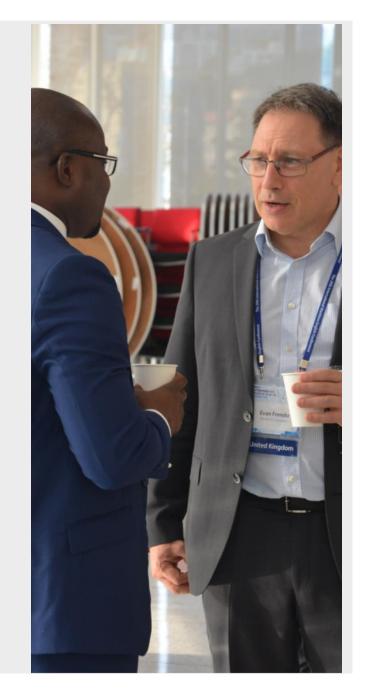
Specialist lexis:

- Everyday lexis with specific meanings in certain contexts
- Technical lexis known and used by experts in the field

Everyday lexis with specific meanings in certain contexts

What was the talk about? Mostly about soap.

Context: maritime communication conference (SOAP = Simple Object Access Protocol)



Technical lexis - an email from the Enron corpus

There are two reason that I do not agree with their interpretation. First, section 6.3 begins with "notwithstanding any other provision of the Plan." This indicates that any other payout methodologies described in other sections of the plan which deal with normal distribution at termination do not apply. Second, the language in section 6.3 stating "a single sum distribution of all of the Participant's deferral accounts" indicates that one payment will be made not a cash payment separate from a share distribution.

Making decisions about what lexis to teach

- Observing learners in action
- Discussions with learners / expert insiders
- Discussions with other trainers
- Intuition

- Frequency analysis
- Keyword analysis

Key terms

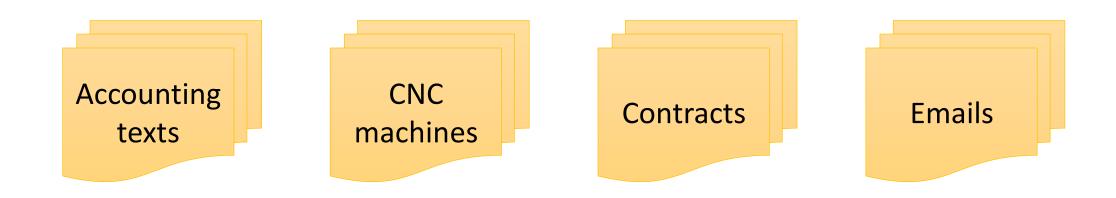
- Corpus
- Concordance
- Frequency
- Collocation
- Cluster / n-gram
- Keyword

What is a corpus?

A collection of texts, spoken or written, held in electronic format.

Two options – build your own, or use one that is available online

Option one - build your own



















Home Resume

Publications *

Software *

Classes

Photo Albums *

Links

Contact

Software



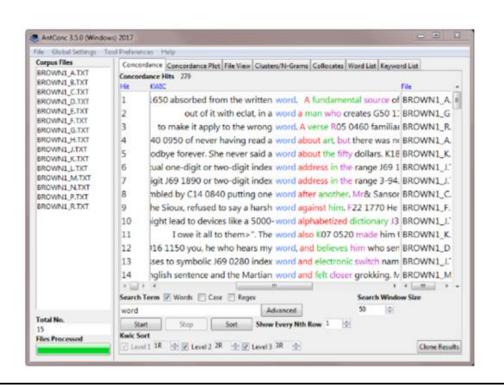
AntConc

A freeware corpus analysis toolkit for concordancing and text analysis.

[AntConc Homepage] [Screenshots] [Help] [License]

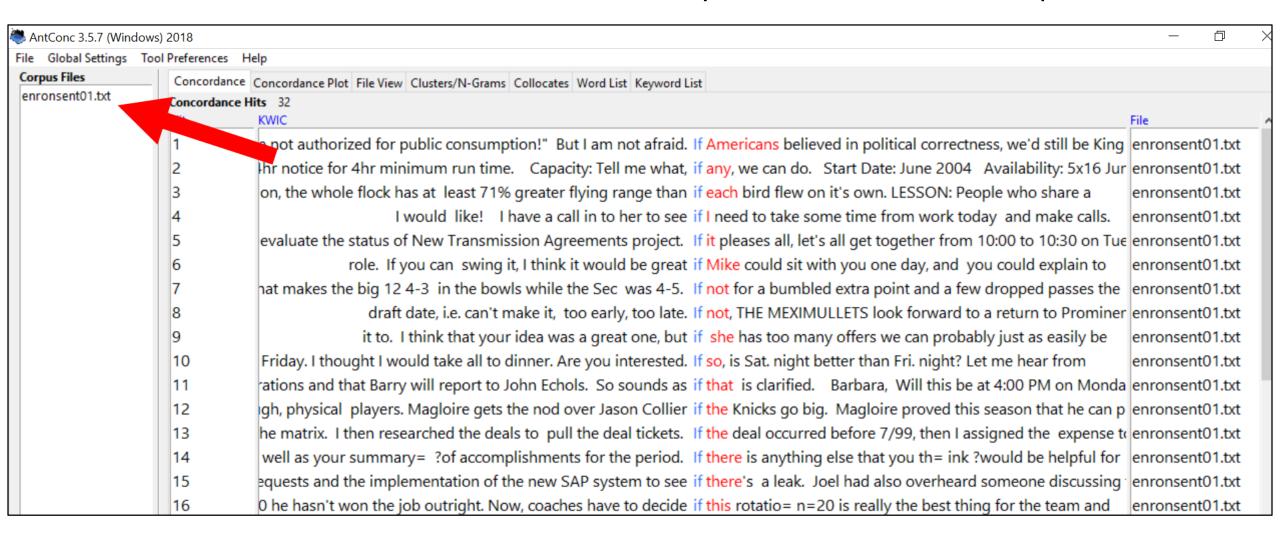
Downloads:

- Windows (3.5.8)
- Windows 64-bit (3.5.8)
- Macintosh OS X 10.6-10.12 (3.5.8)
- Linux 32-bit (3.5.8)
- Linux 64-bit (3.5.8)
- Older versions



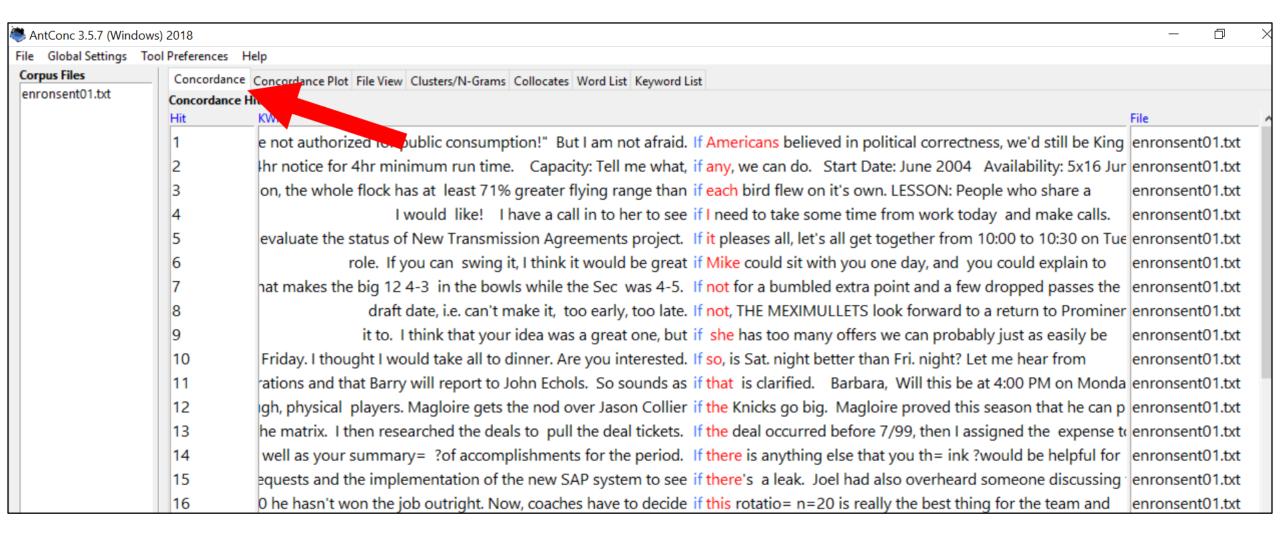
Concordance

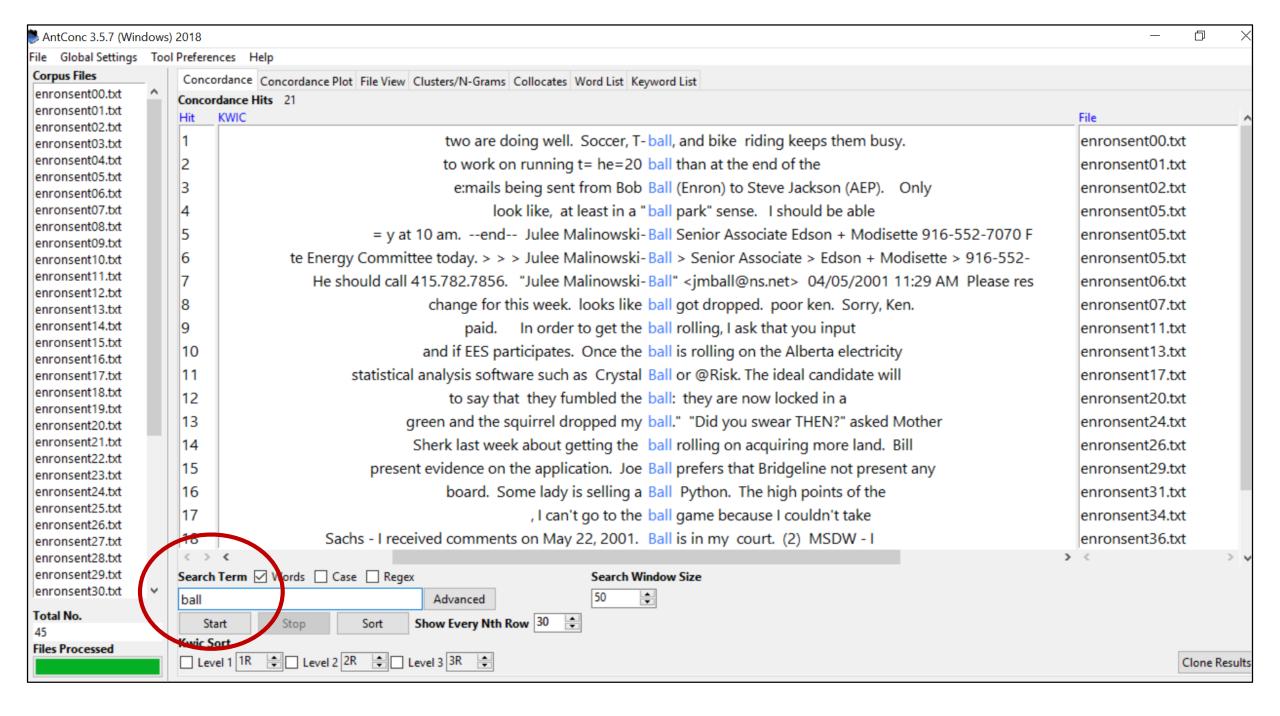
Key word in context (KWIC) Use to find every occurrence of a word or phrase

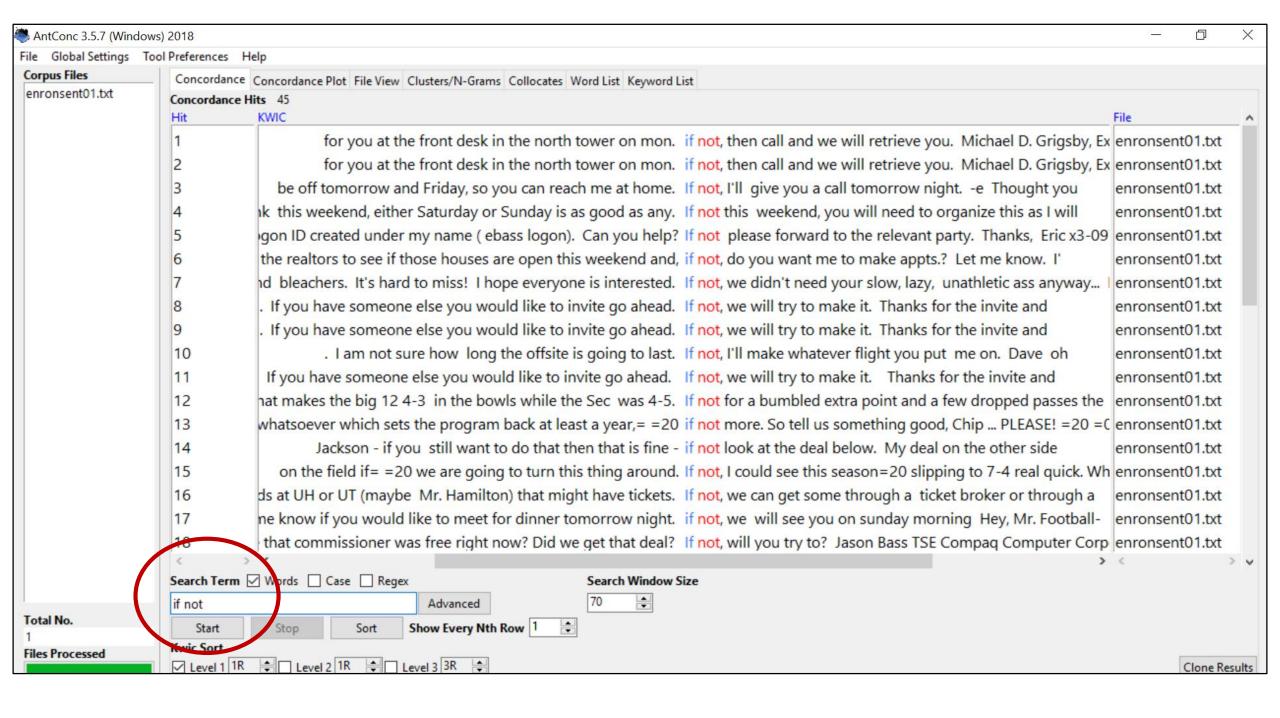


Concordance

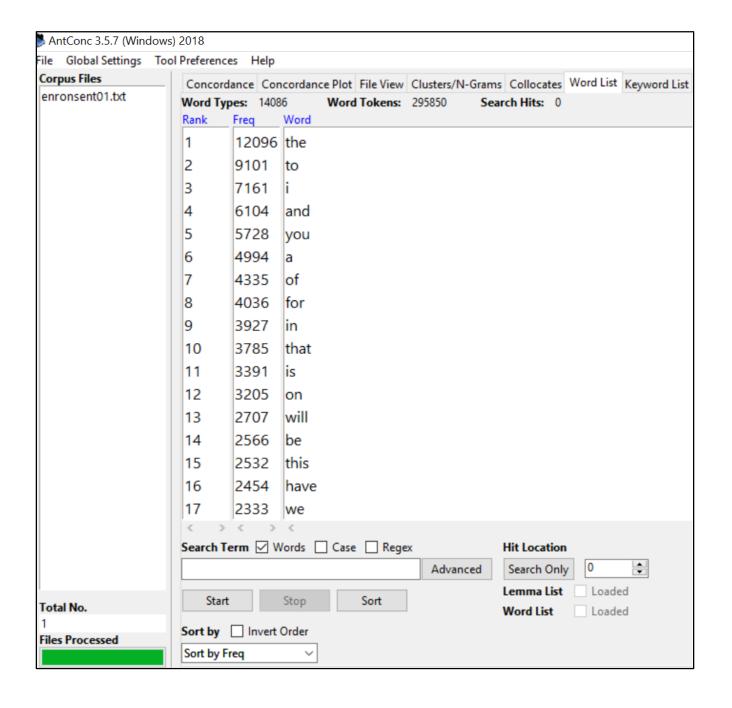
Key word in context (KWIC) Use to find every occurrence of a word or phrase



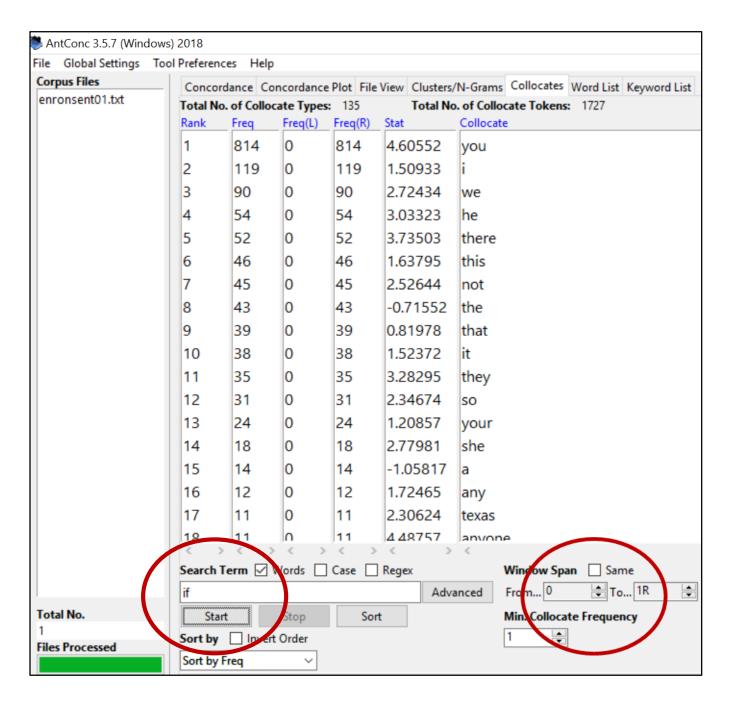




Frequency

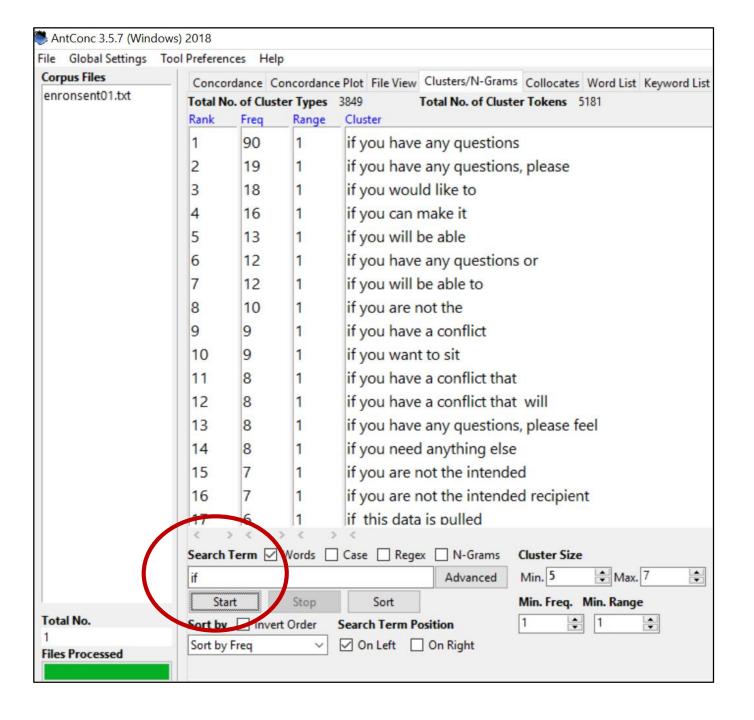


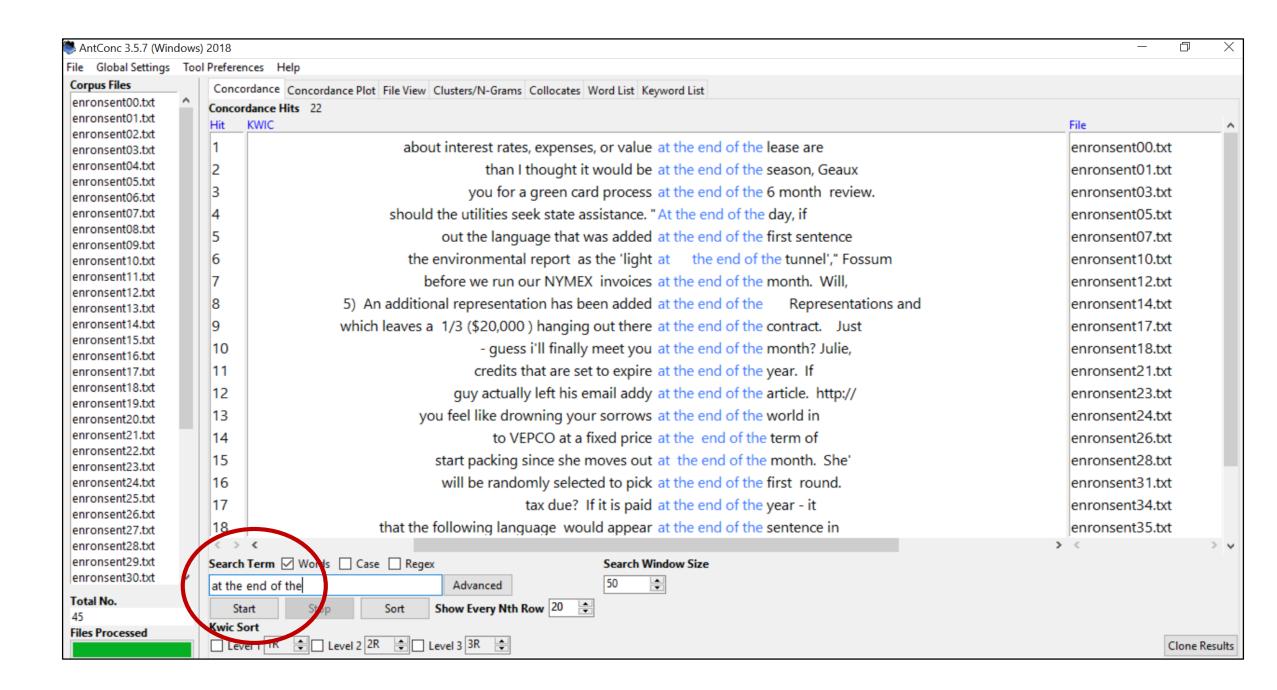
Collocation

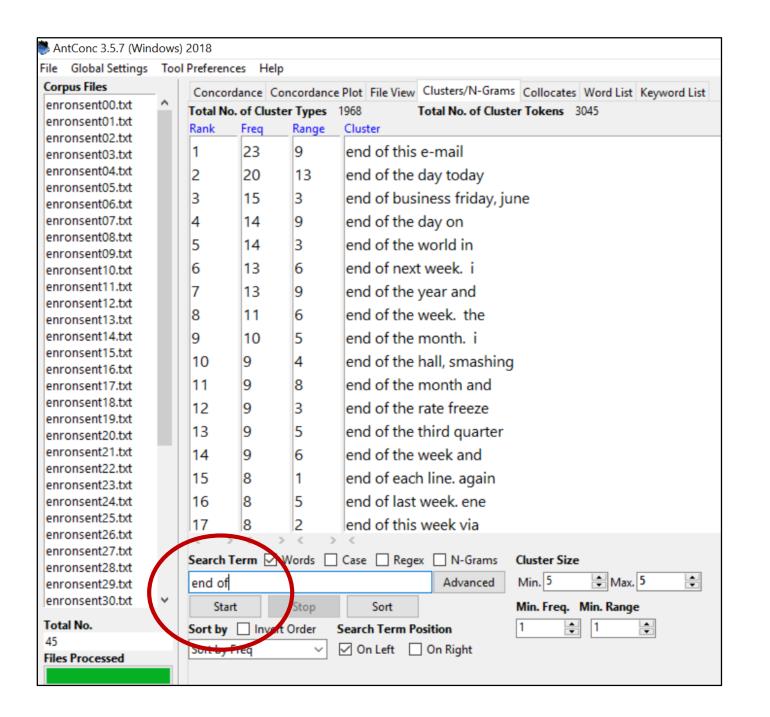


Clusters / N-grams

Word combinations
/ recurrent strings of
2 or more words



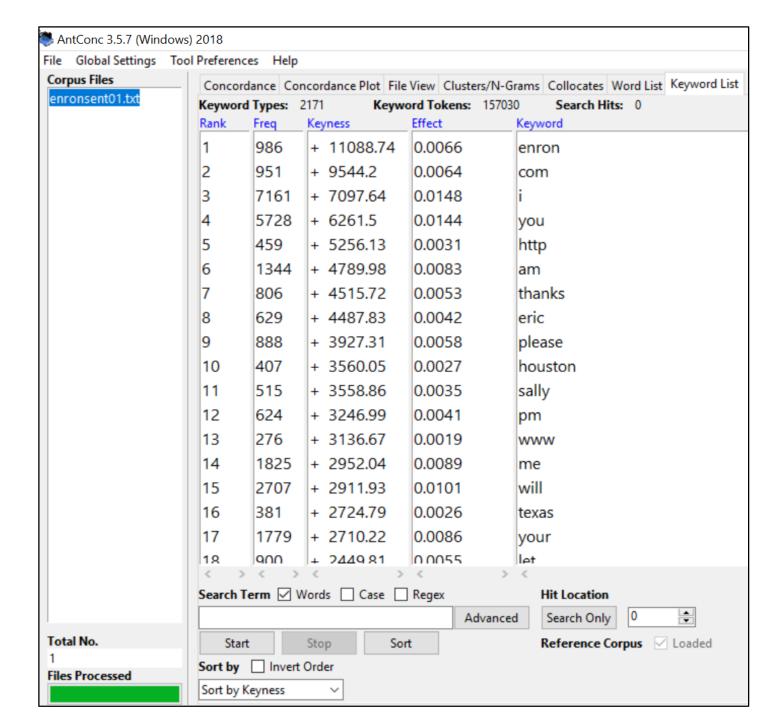




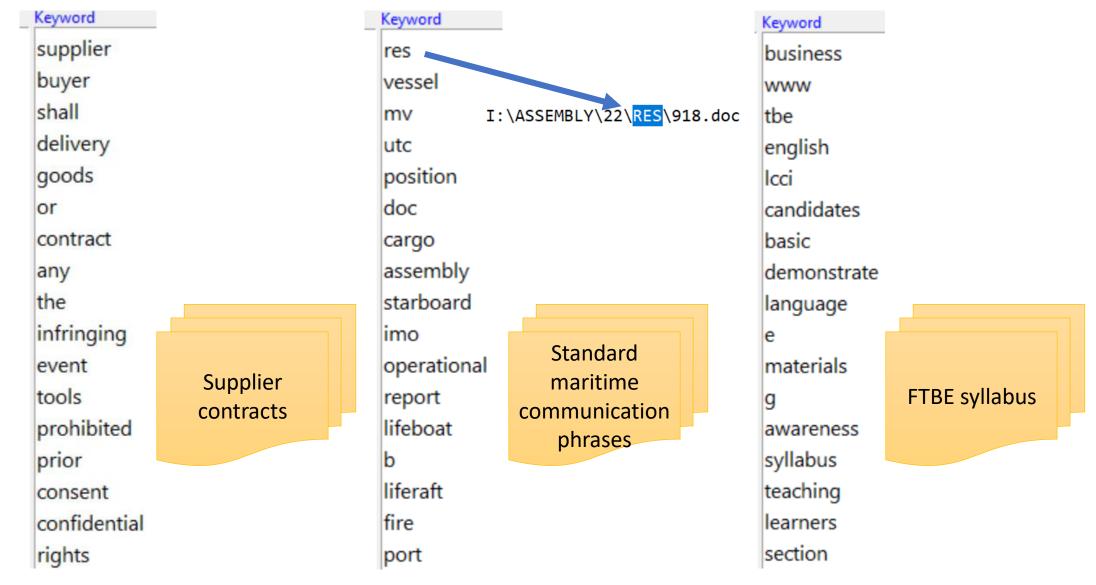
Keyword

Use to find words
 which are
 significantly more
 frequent in a corpus
 (when compared to
 another corpus)

Tells you what the corpus is about



Keyword lists - where do these come from?



Option two - use one that is available online



English-Corpora.org *

home corpora users related resources my account upgrade help

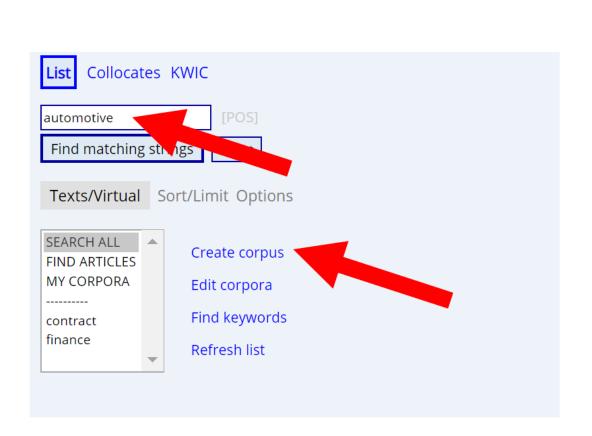
The most widely used online corpora. Overview, search types, looking at variation, corpus-based resources.

The links below are for the online interface. But you can also 🕕 download the corpora for use on your own computer.

Corpus	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus	•	14 billion	6 countries	2017	Web
News on the Web (NOW)	•	8.6 billion+	20 countries	2010-last month	Web: News
Global Web-Based English (GloWbE)	•	1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus	•	1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporal / Ame. English (COCA)	•	560 million	American	1990-2017	Balanced
Corpus of Historical American English (COHA)	•	400 million	American	1810-2009	Balanced
The TV Corpus	•	325 million	6 countries	1950-2018	TV shows
The Movie Corpus	•	200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas	•	100 million	American	2001-2012	TV shows



FREQUENCY



SEARCH



LOGGED IN

ACCOUNT

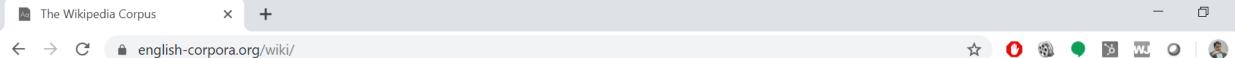
VIRTUAL CORPORA

Create a "virtual corpus" -- essentially your own personalized corpus within WIKI. You can create the corpus either by **keywords in the texts** (e.g. texts with the words *investments*, *basketball*, or *biology*), or **information about the texts** (e.g. date, title, or source), or a combination of keyword and text information.

You can then **edit** your virtual corpora, **Search** within a particular virtual corpus, **compare** the frequency of a word, phrase or grammatical construction in your different virtual corpora, and also create **"keyword lists"** based on the texts in your virtual corpus.

Click on any of the links above for more information.

CONTEXT

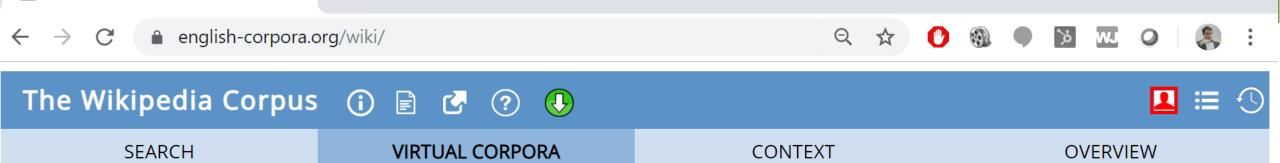




SEARCH VIRTUAL CORPUS CONTEXT OVERVIEW

You can add or remove texts below, and then either save these texts as a new virtual corpus, or else add them to an existing virtual corpus.

SAVE AS: automotive OR ADD TO:SELECT SUBMIT RESET					MIT RESET
HELP	100	# WORDS \$	₹	+ †	TITLE (SEE IN WIKIPEDIA) ‡
1	•	1,919	566	119	Automotive industry
2	•	528	147	20	Aftermarket (automotive)
3	•	2,951	128	134	Automotive design
4	•	1,724	125	38	Spoiler (automotive)
5	•	1,491	116	55	Automotive engineering
6	•	2,786	113	111	Automotive navigation system
7	•	6,613	93	96	Automotive lighting
8	✓	380	68	7	Automotive Hall of Fame
9	✓	170	51	16	Facelift (automotive)
10	✓	1,163	50	27	Diffuser (automotive)
11	✓	2,066	45	37	McLaren Automotive
12	•	526	41	23	Petersen Automotive Museum



MY VIRTUAL CORPORA

The Wikipedia Corpus

HELP	‡	‡	LIST NAME ‡	# ARTICLES ‡	# WORDS \$	FIND KEYWORDS SPECIFIC FREQ	CREATED \$
1			AUTOMOTIVE	100	116,352	NOUN VERB ADJ ADV N+N ADJ+N	0 h
2			CONTRACT	100	179,992	NOUN VERB ADJ ADV N+N ADJ+N	6 h
3			FINANCE	100	140,757	NOUN VERB ADJ ADV N+N ADJ+N	1 d

AUTOMOTIVE [116,352 WORDS, 100 TEXTS] NOUN VERP ADJ ADV N+N ADJ+N

HELP	WORD (CLICK FOR CONTEXT)	FRE	# TEXTS	SPECIFIC FREQ 30 10 TEXTS
1	CAR	946	72	49.3
2	VEHICLE	828	73	115.9
3	PRODUCTION	385	53	11.5
4	COMPANY	351	64	5.2
5	YEAR	330	52	3.3
6	ENGINE	325	48	23.5
7	SYSTEM	308	49	5.7
8	MODEL	273	36	13.8
9	AUTOMOBILE	235	48	114.4
10	DESIGN	205	46	7.8
11	PART	200	55	2.3
12	INDUSTRY	197	38	11.0
13	BATTERY	188	21	51.5
1.4	LAMD	100	5	100.7

AUTOMOTIVE [116,352 WORDS, 100 TEXTS] NOUN VERB ADJ ADV N+N ADJAN

HELP	WORD (CLICK FOR CONTEXT)	FREQ	# TEXIS	FREQ 10 1 TEXTS
1	MODEL YEAR	109	4	222.2
2	TURN SIGNAL	42	2	1,338.3
3	PASSENGER CAR	39	13	153.8
4	MOTOR VEHICLE	35	15	61.9
5	STOP LAMP	28	1	7,950.0
6	SPORT CAR	27	16	48.5
7	NAVIGATION SYSTEM	27	2	147.9
8	PRODUCTION CAR	26	8	241.5
9	AUTOMOBILE INDUSTRY	25	11	410.2
10	AUTO INDUSTRY	24	10	489.2
11	CAR MANUFACTURER	23	15	211.3
12	FOG LAMP	23	2	1,060.0
13	DIESEL ENGINE	22	11	37.5
14	ASSEMBLY PLANT	21	10	188.6

AUTOMOTIVE [116,352 WORDS, 100 TEXTS] NOUN VERB ADJ ADV N+N ADJ+N

HELP	WORD (CLICK FOR CONTEXT)	FREQ	# TEXTS	SPECIFIC FREQ 10 1 TEXTS
1	JOINT VENTURE	82	19	76.9
2	AUTOMOTIVE INDUSTRY	52	22	441.7
3	ELECTRIC VEHICLE	38	15	169.4
4	COMMERCIAL VEHICLE	26	12	169.5
5	NEW CAR	19	10	71.2
6	TOP SPEED	17	8	34.6
7	LIGHT TRUCK	15	8	273.8
8	REAR FOG	15	2	3,359.2
9	ELECTRIC CAR	14	10	88.9
10	NEW MODEL	14	4	40.1
11	DOMESTIC CAR	13	1	3,563.8
12	LIMITED PRODUCTION	12	2	210.8
13	OTHER COUNTRY	11	7	5.8
14	LARGE CAR	11	6	193.5

Notable current relationships include: * AB Volvo and Eicher Motors has a 50-50% joint venture called VE Commercial Vehicles. * Anadolu Group and oint venture called VE Commercial Vehicles. * Anadolu Group and Isuzu have a 50-50% joint venture called Anadolu Isuzu. * Beijing Automotive Group have a 50-50% joint venture called Anadolu Isuzu. * Beijing Automotive Group has a joint venture with Daimler called Beijing Benz, both companies ho Beijing Benz, both companies hold a 50-50% stake. both companies also have a **joint venture** called Beijing Foton Daimler Automobile. BAG also has a j also have a joint venture called Beijing Foton Daimler Automobile. BAG also has a joint venture with Hyundai called Beijing Hyundai, both companies h Hyundai, both companies hold a 50-50% stake. * BMW and Brilliance have a joint venture called BMW Brilliance. BMW owns a 50% stake, Brilliance owi he Shenyang municipal government owns a 9.5% stake. * Chang'an Automobile Group has three **joint ventures**, one with PSA Peugeot Citroen(CAPSA): 35% stake, and Mazda holds a 15% stake. Chang'an and Ford have a joint venture called Chang'an Ford Nanjing. Jiangling and Chang'an have a joint ver Ford have a joint venture called Chang'an Ford Nanjing. Jiangling and Chang'an have a joint venture called Jiangxi Jiangling. * Chery has a joint venture Jiangling and Chang'an have a joint venture called Jiangxi Jiangling. * Chery has a joint venture with Tata Motors called Chery Jaguar Land Rover, both Rover, both companies hold a 50-50% stake. Chery and Israel Corporation has a joint venture called Qoros, both companies hold a 50-50% stake. * Dai and an 85% stake in Master Motors. Daimler and BYD Auto have a joint venture called Denza, both companies hold a 50-50% stake. * Dongfeng Moto

companies hold a 50-50% stake. * Dongfeng Motor Corporation and Nissan have a 50-50% **joint venture** called Venucia, and another 50-50% joint venture.

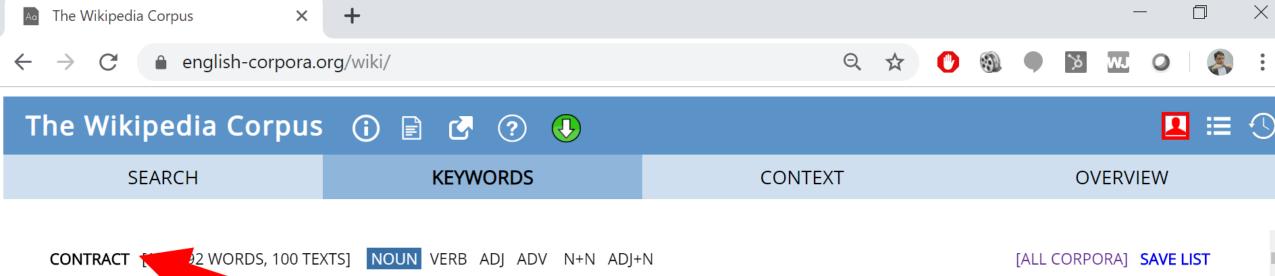
Motor Corporation and Nissan have a 50-50% joint venture called Venucia, and another 50-50% **joint venture** called Dongfeng Motor Company. Dongfe

50-50% joint venture called Dongfeng Motor Company. Dongfeng and PSA Group have a 50-50% **joint venture** called Dongfeng Peugeot-Citroen, a 50-5

Dongfeng and PSA Group have a 50-50% joint venture called Dongfeng Peugeot-Citroen, a 50-50% **joint venture** with Honda called Dongfeng Honda, a joint venture

Dongfeng and PSA Group have a 50-50% joint venture called Dongfeng Peugeot-Citroen, a 50-50% **joint venture** with Honda called Dongfeng Honda, a called Dongfeng Honda, a goint venture with AB Volvo called Dongfeng Nissan-Die Talled Dongfeng Peugeot-Citroen, a 50-50% joint venture with Honda called Dongfeng Honda, a **joint venture** with AB Volvo called Dongfeng Nissan-Die Talled Dongfeng Peugeot-Citroen, a 50-50% joint venture with Honda called Dongfeng Honda, a **joint venture** with AB Volvo called Dongfeng Nissan-Die at least 190-50.

a 49% share in Jiangling Motors. Ford and Navistar International have a 50-50 joint venture called Blue Dimond Truck. Ford and Sollers JSC have a 50-50 joint venture called Ford Sollers. Both Ford and Ko Holding own a 41%



CONTRA	CT 1 32 WORDS, 100 TEXTS] NOUN V	[ALL CORP	ORAJ SAVE LIST			
HELP	WORD (CLICK FO. ONTEXT)	FREQ	# TEXTS	SPECIFIC FREQ 30 10 TEXTS	ALL WIKIPEDIA	EXPECTED
1	CONTRACT	2220	72	110.9	205,763	20.0
2	PARTY	1003	39	12.5	825,277	80.3
3	LAW	579	48	9.5	626,110	60.9
4	TERM	570	45	17.5	335,286	32.6
5	AGREEMENT	354	31	26.3	138,442	13.5
6	CASE	335	48	9.4	367,133	35.7
7	COURT	322	21	6.0	553,606	53.9
8	RIGHT	319	36	15.9	205,935	20.0
9	PERFORMANCE	316	24	8.9	364,029	35.4
10	OBLIGATION	301	26	249.0	12,423	1.2

Creating learning activities

What does it mean to know a word?

Form

spoken

written

word parts

Meaning

form & meaning

concepts & referents

associations

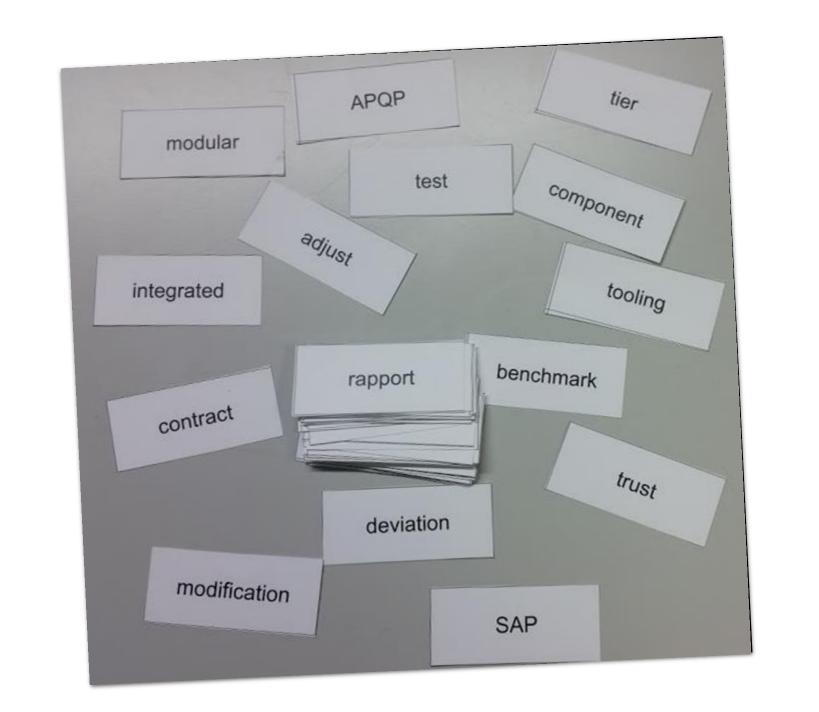
Use

grammatical functions

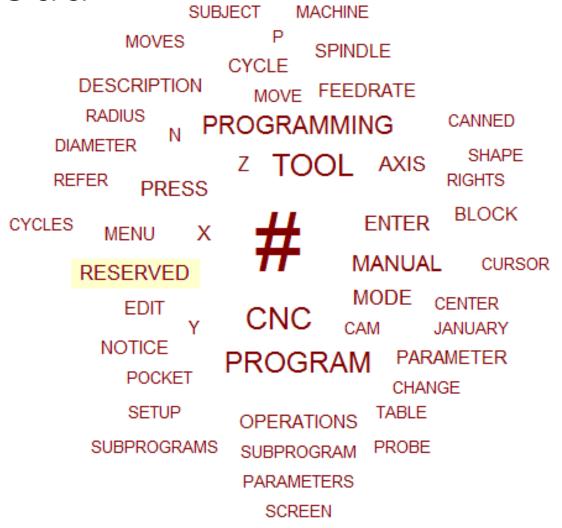
collocations

constraints on use

(Nation, 2001)



Word cloud



Word cloud made from top 50 keywords in a CNC operating manual

Gap-fill (based on concordance lines)

CNC Machines

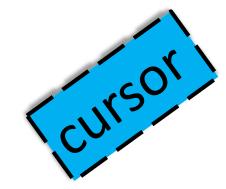
Which word from the word cloud fits in all four lines?

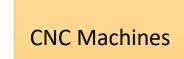
Place the _____ at the node.

Move the _____ to the last block.

Bring the _____ down one line.

Position the _____ over the circle.





Use the words in the box to fill the gaps.

feedrate mode cycle axis radius cycle finding, milling, positioning, facing, roughing, tapping radius specified, 0.25 inch, tool, corner, large, arc feedrate programmed, slow, medium, original, automatic, pass mode manual, edit, draw, auto, incremental, step axis Z, rotary, selected, single, auxiliary, fourth



Use the words in the box to fill the gaps

transactions, engagement, evidence, statement, fraud

- 1 unauthorised, recurring, complex, routine, significant ____
- 2 specific, compilation, group, assurance, audit _____
- 3 further, documentary, additional, appropriate, audit _____
- 4 group, consolidated, summary, income, financial _____
- 5 likely, potential, alleged, possible, suspected _____

Accounting

Use the words in the box to fill the gaps

transactions, engagement, evidence, statement, fraud

- 1 unauthorised, recurring, complex, routine, significant <u>transactions</u>
- 2 specific, compilation, group, assurance, audit engagement
- 3 further, documentary, additional, appropriate, audit evidence
- 4 group, consolidated, summary, income, financial statement
- 5 likely, potential, alleged, possible, suspected <u>fraud</u>



What word from the previous slide fits in all three gaps?

...when undertaking analytical procedures in a limited assurance _____ ...

...agreeing the terms of the compilation _____ ...

...during the course of an audit _____ ...



What word from the previous slide fits in all three gaps?

...when undertaking analytical procedures in a limited assurance engagement...

...agreeing the terms of the compilation engagement...

...during the course of an audit engagement ...

Word formation



Use the words in the box to fill the gaps

acquire, acquired, acquiree, acquirer, acquisition, acquisitive

- 1 They are paying 8 million to _____ the stock.
- 2 We are talking about the net assets _____ by the buyer.
- 3 An _____ may obtain control of an _____ without transferring consideration.
- 4 The _____ date is the date the buyer obtains control of the target.
- 5 The level of _____ activities in the sector has increased.

Use the words in the box to fill the gaps

acquire, acquired, acquiree, acquirer, acquisition, acquisitive

- 1 They are paying 8 million to <u>acquire</u> the stock.
- 2 We are talking about the net assets <u>acquired</u> by the buyer.
- 3 An <u>acquirer</u> may obtain control of an <u>acquiree</u> without transferring consideration.
- 4 The <u>acquisition</u> date is the date the buyer obtains control of the target.
- 5 The level of <u>acquisitive</u> activities in the sector has increased.

Ball – 602 instances in the Enron corpus



On the ball
Get the ball rolling
Keep the ball rolling
Move this ball forward
Ball park
Ball games
Ball in my court
Drop the ball
Play hard ball

Whole ball of wax
A wrecking ball
Have a ball



https://pixabay.com/photos/ball-sport-game-summer-leisure-3343111/

Get the ball rolling



- I thought I'd get the ball rolling and try to get ...
- In order to get the ball rolling, I ask that you ...
- Hopefully this will at least get the ball rolling. It is then up to ...
- ... a draft of the agreement to "get the ball rolling" again. I believe I have included ...
- I had to call today and get the ball rolling because it was ...
- ... and would like to get the ball rolling as soon as possible.
- Anyway, just some thoughts to get the ball rolling. I'm open to just about anything.





You can count on us not dropping the ball.

They really dropped the ball on that one, although I wish I had ...

XXX dropped the ball, unfortunately.

Sorry if I dropped the ball

Someone REALLY **dropped the ball** here!

If we dropped the ball on this, I apologize.

While I am not sure who **dropped the ball** on this item, the fact remains ...

I owe you an apology for dropping the ball





You can count on us not dropping the ball.

They really dropped the ball on that one, although I wish I had ...

XXX dropped the ball, unfortunately.

Sorry if I dropped the ball

Someone REALLY dropped the ball here!

If we dropped the ball on this, I apologize.

While I am not sure who **dropped the ball** on this item, the fact remains ...

I owe you an apology for dropping the ball.

Drop the ball



Just wanted to let you know that I've received all the ... stuff and that you can count on me to completely **drop** the ball.

Kidding! I'll take care of it.

Outline

- What is specialist lexis?
- Making decisions about what lexis to teach
 - Corpus analysis tools
- Creating learning activities

