

Corpus Linguistics as a methodological tool for ESP teaching

Dr. Natalia Judith Laso

Tashkent, 14-15 January 2016



UNIVERSITAT DE
BARCELONA



Grup de Recerca
en Lexicologia i
Lingüística de Corpus

Corpus Linguistics as a methodological tool for ESP teaching

➤ **Corpus:**

- collection of texts & language samples
- it provides **evidence of real language in use**
(more reliable than NS's intuition)
- A way of collecting and storing data, which can be accessed by means of corpus query programmes (WordSmith Tools, AntConc, The Sketch Engine, etc.)
- A corpus tells us **what language is like**
(**descriptive approach**)

Corpus Linguistics as a methodological tool for ESP teaching

- A corpus is **a more reliable guide to language use than NS intuition is**. Cook (2001) points out that although a NS has experience of very much more language than is contained in even the largest of corpus, much of that experience remains hidden from introspection.

Corpus Linguistics as a methodological tool for ESP teaching

- When a learner asks why a particular phrasing is preferred in a particular context to another, a language teacher usually replies that “ it just sounds better” (source of irritation to learners! 🤔)
- **MAIN USES:** research/teaching purposes; translation; contrastive studies; discourse analysis

Corpus Linguistics as a methodological tool for ESP teaching

PROS	CONS
<p>RELIABILITY: results are NOT based on human intuition, which is a poor guide to collocation, phraseology, prosody and phraseology.</p> <p>Frequency counts: lexis, grammar</p>	<p>It does not give information about whether something is possible or not, only whether it is frequent or not (Sinclair 1991; Biber <i>et al.</i> 1998; McEnery & Wilson 2001; Hunston 2002)</p>
<p>It provides evidence for collocations and phraseology: access to plenty of examples of real language in use.</p> <p>It is stored electronically: it can be studied non-linearly, and both quantitatively & qualitatively</p>	<p>It may claim to be representative, but all attempts to draw generalisations from a corpus are in fact extrapolations (Hunston 2002).</p> <p>WATCH OUT! A corpus is always a SAMPLE of language.</p>

Corpus Linguistics as a methodological tool for ESP teaching

➤ TYPOLOGY:

- ❖ **Specialised corpus** (of a variety of language/domain): **representative of a given type of text & used to investigate a particular type of language** (e.g., *Health Science Corpus*)
- ❖ **General corpus**: a corpus of **texts of many types; much larger than a specialised one**. It may be used to produce reference materials for language learning or translation, and it's **often used as a baseline for comparison against more specialised corpus** (i.e., as a *reference corpus*) (e.g., *British National Corpus; Bank of English*)

Corpus Linguistics as a methodological tool for ESP teaching

- ❖ **Comparable corpus:** two or more corpora in different languages (Eng. & Sp.) or in different varieties of a language (BrE vs. AmE). They are designed along the same lines (e.g., same size, same typology of texts, etc.)
- ❖ **Parallel (translated) corpus:** two or more corpora in different languages (Eng. & Sp.), each containing texts that have been translated from one language to another.

Corpus Linguistics as a methodological tool for ESP teaching

- ❖ **Learner corpus:** a collection of texts produced by learners of a language; aims at identifying how the language of NNS differs from the language of NS. (e.g. *The International Corpus of Learner English*, the [VESPA project](#))
- ❖ **Diachronic corpus:** a corpus of texts from different periods of time, so as to trace the development of linguistic aspects throughout different periods of time (e.g., *Helsinki Corpus*, a 1.5 million word corpus which consists of English texts from 700 to 1700)

Corpus Linguistics as a methodological tool for ESP teaching

- Issues to bear in mind in corpus design:
 - **Corpus type:** general/specialised corpus?
 - **Content**
 - **Parts & type of files** (sections & subsections?/
whole texts or extracts from texts?)
 - **Size** (manageable data!)
 - **Balance & Representativeness**
 - **Permanence**

Corpus Linguistics as a methodological tool for ESP teaching

“It is no exaggeration to say that **corpora**, and the study of corpora, **have revolutionised the study of language, and of the applications of language**, over the last few decades. The improved accessibility of computers has changed corpus study from a subject for specialists only to **something that is open to all.**” (Hunston 2002, p. 1)

Corpus Linguistics as a methodological tool for ESP teaching

What does processing data provide?

- **Frequency**
- **Collocations**
- **Phraseology**

Corpus Linguistics as a methodological tool for ESP teaching

The **words** in a corpus can be **arranged in order of their frequency** in that corpus.

Possible applications: comparing frequency lists from different corpora

Example 1:

Below are the top 25 words in a corpus of politics dissertations compared with a corpus of materials science dissertations and with the *Bank of English* corpus (University of Birmingham)

Corpus Linguistics as a methodological tool for ESP teaching

Table 1.1. Word frequency comparisons across corpora

	General corpus	Materials science	Politics
1	THE	THE	THE
2	OF	OF	OF
3	TO	AND	TO
4	AND	IN	AND
5	A	TO	IN
6	IN	A	A
7	THAT	IS	THAT
8	S	THIS	IS
9	IS	P	AS
10	IT	THAT	WAS
11	FOR	FOR	FOR
12	I	BE	IT
13	WAS	AS	THIS
14	ON	HEAD	P
15	HE	ARE	ON
16	WITH	WITH	BE
17	AS	IT	BY
18	YOU	BY	WHICH
19	BE	ON	S
20	AT	WAS	NOT
21	BY	AT	INTERNATIONAL
22	BUT	WHICH	WITH
23	HAVE	FROM	AN
24	ARE	FIGURE	QUOTE
25	HIS	AN	ARE

Corpus Linguistics as a methodological tool for ESP teaching

26	FROM	NOT	FROM
27	THEY	HAS	WERE
28	THIS	WERE	<u>POLICY</u>
29	NOT	CAN	<u>STATES</u>
30	HAD	THESE	BUT
31	HAS	BEEN	<u>STATE</u>
32	AN	HAVE	<u>WOULD</u>
33	WE	OR	OR
34	N'T	<u>SURFACE</u>	ITS
35	OR	USED	MAZZINI
36	<u>SAID</u>	C	THEIR
37	ONE	ENERGY	HEAD
38	THERE	<u>TEMPERATURE</u>	AT
39	WILL	ALSO	HAD
40	THEIR	WILL	HAVE
41	WHICH	CONTRAST	MORE
42	SHE	TWO	BRITAIN
43	WERE	FIELD	THEY

Corpus Linguistics as a methodological tool for ESP teaching

Another possible application is to look at the **frequency of given words**, compared across corpora

Example 2:

Below are the frequency counts of *must*, *have to*, *incredibly* and *surprisingly* across corpora (per million words) (Hunston 2002:6)

Corpus Linguistics as a methodological tool for ESP teaching

	Books	<i>Times</i>	Spoken
<i>must</i>	683	460	363
<i>have to</i>	419	371	802
Total	1102	831	1165
<i>incredibly</i>	8	10	15
<i>surprisingly</i>	25	29	4
Total	33	39	19

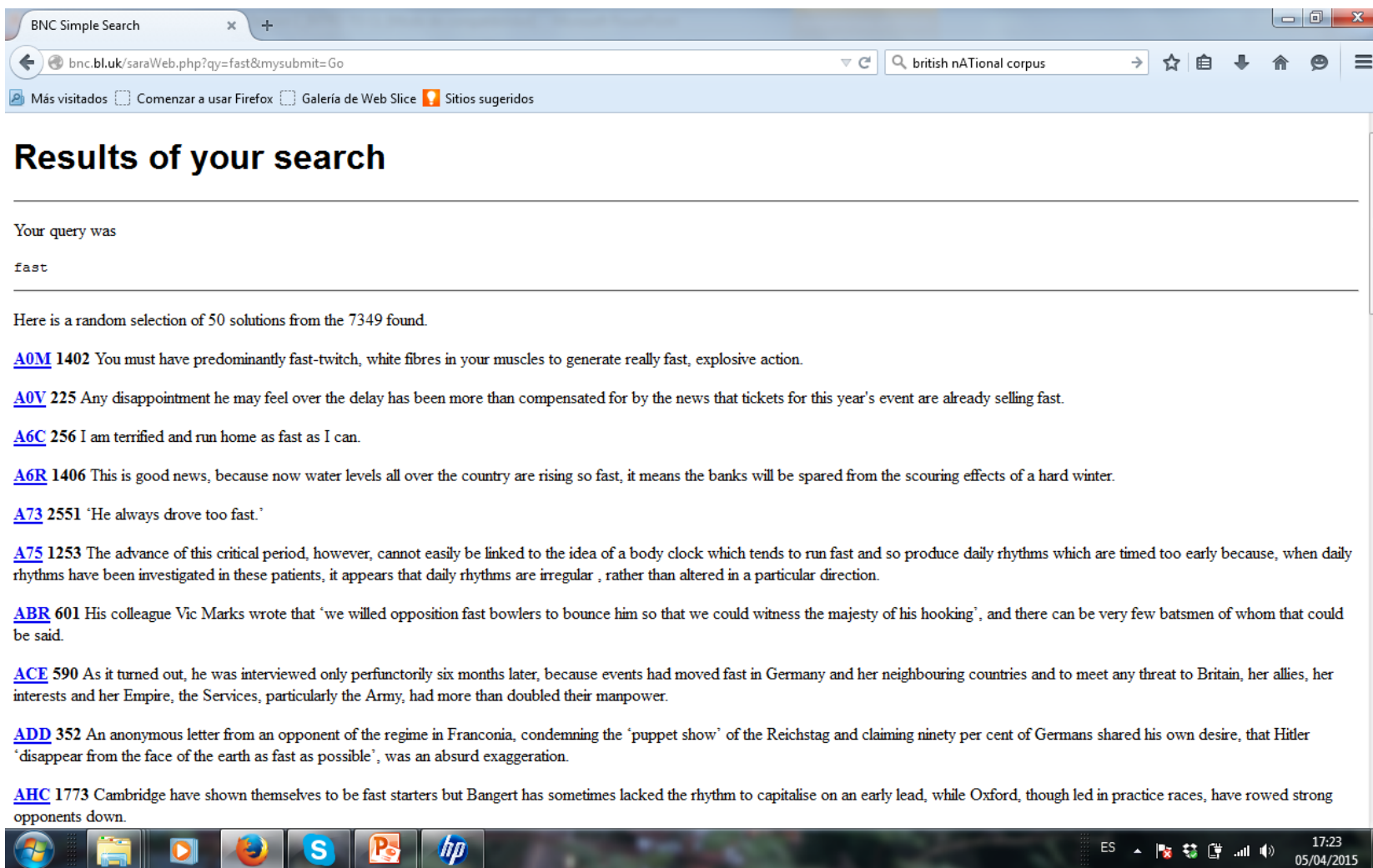
Corpus Linguistics as a methodological tool for ESP teaching

Another approach is to compare the **frequency of use of two lemmas** in a corpus

Example 3:

Below are two screenshots from the *BNC*, each of which shows the frequency counts of the lemmas *fast* and *quick*, respectively

Corpus Linguistics as a methodological tool for ESP teaching

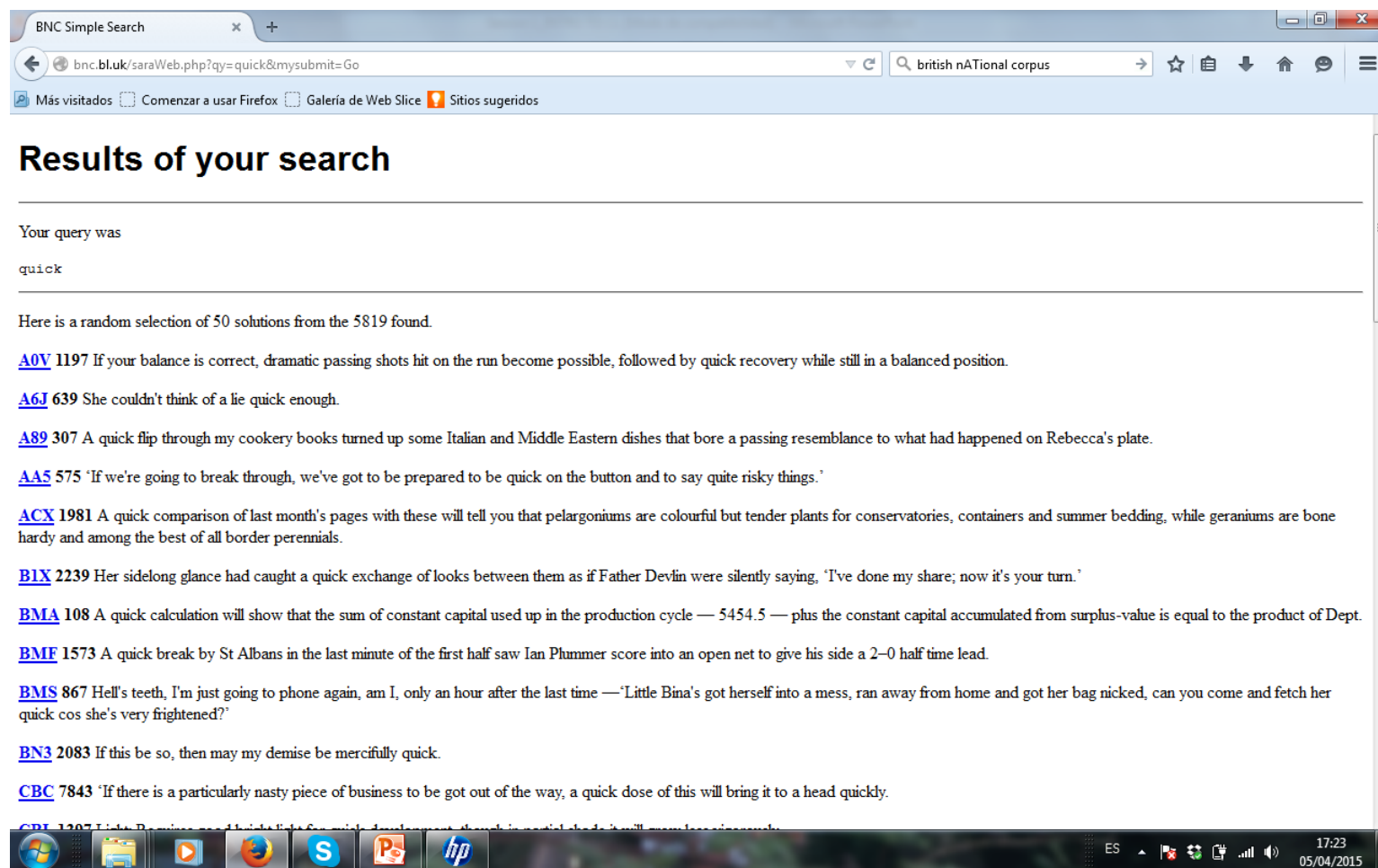


The screenshot shows a web browser window titled "BNC Simple Search". The address bar contains the URL "bnc.bl.uk/saraWeb.php?qy=fast&mysubmit=Go". The search bar shows the query "fast". The page title is "Results of your search". Below the title, it says "Your query was" followed by "fast". A message states: "Here is a random selection of 50 solutions from the 7349 found." The results are listed as follows:

- [A0M](#) 1402 You must have predominantly fast-twitch, white fibres in your muscles to generate really fast, explosive action.
- [A0V](#) 225 Any disappointment he may feel over the delay has been more than compensated for by the news that tickets for this year's event are already selling fast.
- [A6C](#) 256 I am terrified and run home as fast as I can.
- [A6R](#) 1406 This is good news, because now water levels all over the country are rising so fast, it means the banks will be spared from the scouring effects of a hard winter.
- [A73](#) 2551 'He always drove too fast.'
- [A75](#) 1253 The advance of this critical period, however, cannot easily be linked to the idea of a body clock which tends to run fast and so produce daily rhythms which are timed too early because, when daily rhythms have been investigated in these patients, it appears that daily rhythms are irregular, rather than altered in a particular direction.
- [ABR](#) 601 His colleague Vic Marks wrote that 'we willed opposition fast bowlers to bounce him so that we could witness the majesty of his hooking', and there can be very few batsmen of whom that could be said.
- [ACE](#) 590 As it turned out, he was interviewed only perfunctorily six months later, because events had moved fast in Germany and her neighbouring countries and to meet any threat to Britain, her allies, her interests and her Empire, the Services, particularly the Army, had more than doubled their manpower.
- [ADD](#) 352 An anonymous letter from an opponent of the regime in Franconia, condemning the 'puppet show' of the Reichstag and claiming ninety per cent of Germans shared his own desire, that Hitler 'disappear from the face of the earth as fast as possible', was an absurd exaggeration.
- [AHC](#) 1773 Cambridge have shown themselves to be fast starters but Bangert has sometimes lacked the rhythm to capitalise on an early lead, while Oxford, though led in practice races, have rowed strong opponents down.

The browser's taskbar at the bottom shows various icons including Internet Explorer, Firefox, and several application icons. The system clock in the bottom right corner displays the time 17:23 and the date 05/04/2015.

Corpus Linguistics as a methodological tool for ESP teaching



The screenshot shows a web browser window titled "BNC Simple Search". The address bar contains the URL "bnc.bl.uk/saraWeb.php?qy=quick&mysubmit=Go". The search bar on the right contains the text "british nAtional corpus". Below the search bar, there are several search results listed, each with a reference code and a sentence snippet. The results are as follows:

- A0V 1197** If your balance is correct, dramatic passing shots hit on the run become possible, followed by quick recovery while still in a balanced position.
- A6J 639** She couldn't think of a lie quick enough.
- A89 307** A quick flip through my cookery books turned up some Italian and Middle Eastern dishes that bore a passing resemblance to what had happened on Rebecca's plate.
- AA5 575** 'If we're going to break through, we've got to be prepared to be quick on the button and to say quite risky things.'
- ACX 1981** A quick comparison of last month's pages with these will tell you that pelargoniums are colourful but tender plants for conservatories, containers and summer bedding, while geraniums are bone hardy and among the best of all border perennials.
- BLX 2239** Her sidelong glance had caught a quick exchange of looks between them as if Father Devlin were silently saying, 'I've done my share; now it's your turn.'
- BMA 108** A quick calculation will show that the sum of constant capital used up in the production cycle — 5454.5 — plus the constant capital accumulated from surplus-value is equal to the product of Dept.
- BMF 1573** A quick break by St Albans in the last minute of the first half saw Ian Plummer score into an open net to give his side a 2-0 half time lead.
- BMS 867** Hell's teeth, I'm just going to phone again, am I, only an hour after the last time — 'Little Bina's got herself into a mess, ran away from home and got her bag nicked, can you come and fetch her quick cos she's very frightened?'
- BN3 2083** If this be so, then may my demise be mercifully quick.
- CBC 7843** 'If there is a particularly nasty piece of business to be got out of the way, a quick dose of this will bring it to a head quickly.'
- GBL 1207** Little Demise and his light for quick development, though in special case it will grow less in search.

The browser window also shows a taskbar at the bottom with various application icons and a system clock indicating 17:23 on 05/04/2015.

Corpus Linguistics as a methodological tool for ESP teaching

- Let's have a look at some *BNC* corpus data provided by The Sketch Engine ([©Lexical Computing](#)):
 - a) Which one is more frequently used?
 - b) What kind of contexts allow for both?
 - c) What contexts do show a preference for *fast* or *quick*?

Corpus Linguistics as a methodological tool for ESP teaching

BNC Simple Search WS Diff for fast and quick

https://the.sketchengine.co.uk/bonito/run.cgi/wsdiff?corpname=preloaded%2Fbnc2&reload=&lemma=fast&lpos=-j&diff_by=lemm sketch engine

Más visitados Comenzar a usar Firefox Galería de Web Slice Sitios sugeridos

user: Nataka Laso corpus: British National Corpus

Search in British National Corpus

Concordance
Word List
Word Sketch
Thesaurus
Find X
Sketch-Diff
Corpus Info
?

Change options

Menu position

fast/quick

British National Corpus freqs = 3,065 | 5,552

fast	6.0	4.0	2.0	0	-2.0	-4.0	-6.0	quick
and/or	790	954	2.5	1.9				
furious	30	0	9.0	--				
overnight	16	0	8.4	--				
flowing	12	0	8.3	--				
moving	14	0	8.2	--				
expanding	7	0	7.2	--				
forward	8	0	6.9	--				
afloat	4	0	6.9	--				
hard	72	0	6.9	--				
pakistani	4	0	6.8	--				
loose	12	0	6.8	--				
growing	23	0	6.7	--				
slow	22	5	6.9	4.7				
efficient	15	19	6.6	6.9				
convenient	4	9	5.5	6.6				
simple	7	42	3.9	6.5				
easy	8	112	3.9	7.7				
tasty	0	4	--	6.6				
cheap	0	15	--	6.7				
decisive	0	7	--	6.7				
nimble	0	4	--	6.9				
sideways	0	5	--	7.1				
modifier	531	708	0.8	0.6				
blindingly	4	0	7.9	--				
incredibly	6	0	7.4	--				
both	7	0	3.8	--				
extremely	24	5	6.9	4.6				
really	19	11	4.7	3.9				
too	108	86	6.7	6.4				
very	140	173	5.8	6.1				
particularly	4	6	3.6	4.1				
as	65	101	6.2	6.8				
that	4	8	5.4	6.3				
so	54	110	5.0	6.1				
fairly	5	12	4.8	6.1				
quite	8	21	3.5	4.9				
also	5	15	1.1	2.7				
pretty	4	20	4.9	7.2				
less	0	4	--	2.9				
usually	0	4	--	3.6				
always	0	11	--	3.8				
right	0	4	--	4.6				
bit	0	7	--	4.9				
reasonably	0	4	--	5.5				
modifies	2,004	2,967	3.6	3.2				
reactor	62	0	8.8	--				
lane	63	0	8.3	--				
bowling	23	0	7.9	--				
forward	13	0	7.3	--				
bacillus	9	0	7.1	--				
pace	28	0	6.9	--				
track	49	0	6.8	--				
breeder	10	0	6.8	--				
outlet	14	0	6.6	--				
bowler	118	7	9.9	5.5				
buck	11	13	7.0	6.8				
nod	0	11	--	6.5				
kiss	0	16	--	6.5				
smile	0	42	--	6.6				
wit	0	15	--	6.6				
getaway	0	10	--	6.7				
flick	0	11	--	6.7				
check	0	26	--	6.8				
thinking	0	35	--	6.8				
wash	0	18	--	7.1				
throw-in	0	18	--	7.6				

Feedback

ES 17:28 05/04/2015

Corpus Linguistics as a methodological tool for ESP teaching

Another example of how corpus data can be exploited is the calculation and, most importantly, analysis of **collocations**

What is a **collocation**?: **Statistical tendency of words to co-occur**

Collocations can indicate pairs of **lexical items** (e.g., *shed light; ride a horse; brew tea; critically ill*), or the association between a **lexical word + its frequent grammatical environment (colligation)**, such as *result in; his head; head of the department; a ... (bunch, group, type, part, etc.) of; the ... (woman, man, problem, etc.) in*

Corpus Linguistics as a methodological tool for ESP teaching

Concord																			
File Edit View Compute Settings Windows Help																			
N	Word	With	Relation	Texts	Total	total Left	total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5	
1	RESULTS	results	0.000	11	2,350	9	9	4	1	2	2	0	2,332	0	2	2	1	4	
2	THE	results	0.000	11	1,279	684	595	62	66	61	82	413	0	5	198	145	136	111	
3	RESULT	result	0.000	9	843	0	0	0	0	0	0	0	843	0	0	0	0	0	
4	IN	results	0.000	9	747	155	592	36	44	43	27	5	0	378	50	62	57	45	
5	OF	results	0.000	11	739	266	473	53	65	66	72	10	0	149	1	63	152	108	
6	THAT	results	0.000	8	736	68	668	13	8	8	20	19	0	6	444	80	80	58	
7	THESE	results	0.000	7	578	542	36	5	6	5	12	514	0	1	20	8	5	2	
8	OF	result	0.000	8	526	86	440	35	28	16	7	0	0	235	2	46	105	52	
9	RESULTING	resulting	0.000	10	459	0	0	0	0	0	0	0	459	0	0	0	0	0	
10	THE	result	0.000	8	457	209	248	33	33	11	18	114	0	9	112	30	48	49	
11	IN	result	0.000	9	406	69	337	25	23	12	9	0	0	253	3	24	33	24	
12	IN	resulted	0.000	8	395	32	363	14	10	4	4	0	0	298	1	3	28	33	
13	THE	resulting	0.000	10	372	285	87	30	34	13	27	181	0	0	41	2	21	23	
14	RESULTED	resulted	0.000	8	327	0	0	0	0	0	0	0	327	0	0	0	0	0	
15	A	results	0.000	8	320	48	272	22	14	9	2	1	0	3	177	33	33	26	
16	A	result	0.000	8	317	195	122	6	6	3	24	156	0	0	76	12	14	20	
17	WITH	results	0.000	8	310	135	175	10	24	29	59	13	0	22	31	73	25	24	
18	AND	results	0.000	8	283	156	127	24	22	35	57	18	0	15	6	23	41	42	
19	OUR	results	0.000	8	262	231	31	4	5	1	14	207	0	0	18	6	7	0	
20	IN	resulting	0.000	10	256	37	219	10	3	14	10	0	0	175	3	7	20	14	
21	NOT	results	0.000	7	251	22	229	5	11	5	0	1	0	187	25	3	6	8	
22	SHOWN	results	0.000	6	236	6	230	1	2	3	0	0	0	16	199	8	0	7	
23	OF	resulting	0.000	9	208	96	112	14	15	21	45	1	0	0	6	38	48	20	
24	THIS	result	0.000	6	206	199	7	8	12	13	15	151	0	0	6	0	1	0	
25	OF	resulted	0.000	7	202	83	119	25	22	18	18	0	0	0	0	18	60	41	
26	SUGGEST	results	0.000	5	192	0	192	0	0	0	0	0	0	148	21	8	8	7	
27	ARE	results	0.000	10	189	11	178	3	4	2	2	0	0	116	9	8	24	21	
28	THAT	result	0.000	8	187	77	110	13	15	14	9	26	0	11	58	14	12	15	
29	TO	results	0.000	9	182	100	82	25	23	21	27	4	0	11	7	21	27	16	
30	FROM	results	0.000	8	176	24	152	2	6	4	12	0	0	115	13	7	4	13	
31	WERE	results	0.000	8	147	11	136	5	4	1	0	1	0	103	5	8	8	12	
32	THE	resulted	0.000	8	146	65	81	12	27	19	7	0	0	0	51	2	4	24	
33	INDICATE	results	0.000	5	141	1	140	1	0	0	0	0	0	108	9	10	5	8	

concordance collocates plot patterns clusters filenames follow up source text notes

835 Type-in e controls, in the absence of added cytokines (Fig. 1, A and C). The number of colonies that grew in the absence of added growth factors is relatively low compared with the number of colonies in the presence of factors. The number of cells pl

Correu electrònic ... D:\TESIS\CORPUS FL... Microsoft PowerPoi... WordSmith Tools Concord ES 10:12

Corpus Linguistics as a methodological tool for ESP teaching

Concord

FileEditViewComputeSettingsWindowsHelp

N	Concordance	Set	Tag	Word #	Sent %	Sent	Para %	Para	lead %	lead	Sec %	Sec	File
1	polymorphism in MHC genes Table 4. Summary of polymorphism results. G34 Table 1. The strains used in this study Table 2. Levels			1,013	10630%		031%				031%		Tables.txt
2	in activated MAP kinase strains Table 9. mpk-1 null mutations result in maternal effect larval lethality. G32 Table 1. Progeny ratios			930	9778%		029%				029%		Tables.txt
3	, ICAM-1, and VCAM-1. JCI26 Table 1. Cholecystokinin-8 does not result in conditioned taste aversion. JCI27 Table I Lung Weight,			2,578	24556%		030%				030%		Tables.txt
4	lines. G38 Table 1. Conflicting site patterns Table 2. The results of the test of equal age of nonsynonymous and			1,153	12520%		036%				036%		Tables.txt
5	cultures (Fig. 5; P < 0.05). In contrast, TIMP-3 overexpression resulted in a significant dose dependent increase in the BrdU index			189,000,703	35%		059%				059%		Results.txt
6	with uninfected controls (Fig. 5). TIMP-2 overexpression, however, resulted in a significant dose-dependent decrease in the			188,968,703	21%		059%				059%		Results.txt
7	effect of RAD66 infection or addition of 10 μ M Ro-31-9370 (results not shown) on cell number was observed at any time point			188,725,692	47%		059%				059%		Results.txt
8	chronic, Table 1), 9 contained EBV DNA demonstrable by ISH. The results from 2 representative specimens are shown in Figure 3. The			189,946,756	25%		059%				059%		Results.txt
9	mechanisms, we examined whether RADTIMP-3 infection resulted in induction of WAF1, a protein associated with the cell			189,390,727	58%		059%				059%		Results.txt
10	nuclei and condensed chromatin (n = 3; Fig. 6, C and D). Similar results were obtained with the terminal dUTP nick end labeling			189,131,714	19%		059%				059%		Results.txt
11	= 12). RAD66 infection did not inhibit chemotaxis (102 \pm 7%). These results indicate that inhibition of chemotaxis contributed in part to			188,651,688	16%		059%				059%		Results.txt
12	complications occurring at weeks 8 and 3 of HAART, respectively, resulted in a transient and reversible increase in susceptibility of			188,002,654	76%		058%				058%		Results.txt
13	apoptosis (12, 17, 20), we investigated whether HAART would result in a modification of the level of susceptibility of peripheral			187,807,650	76%		058%				058%		Results.txt
14	experiments that demonstrate that effective antiretroviral therapy results in a rapid decrease in tissue apoptosis, we investigated the			187,322,633	37%		058%				058%		Results.txt
15	inhibited invasion per se, overexpression of TIMP-1, -2, or -3 resulted in an additional profound dose-dependent inhibition (Fig. 2			188,497,678	71%		059%				059%		Results.txt
16	from uninfected or RAD66-infected cultures. RADTIMP-3 infection resulted in dose-dependent induction of TIMP-3 both in the total cell			188,295,669	14%		058%				058%		Results.txt
17	in the levels of peripheral CD8 T cells (Fig. 6). Altogether, these results indicate that circulating peripheral T cells are in an			188,077,658	12%		058%				058%		Results.txt
18	nor CMV DNA was detectable in any tissue. When these results were correlated with the H&E data described in Table 1, the			190,074,768	12%		059%				059%		Results.txt
19	accessory cells. However, our results are consistent with previous results using primary murine fetal liver cells infected with retrovirus			194,832,974	45%		050%				050%		Results.txt
20	to growth factors produced by accessory cells. However, our results are consistent with previous results using primary murine			194,827,972	20%		050%				050%		Results.txt
21	after (day 7) treatment of HIV-infected subjects with G-CSF. As a result of G-CSF treatment, PMN shifted from a growth-promoting to			194,305,943	20%		050%				050%		Results.txt
22	study period (P = not significant [ns] by ANOVA). Because the results obtained with the two different vehicle preparations were			196,299,033	25%		051%				051%		Results.txt
23	, the biological effects observed in this mouse model are the result of chronic low level expression of activated hc. Hematological			195,302,985	52%		051%				051%		Results.txt
24	that this activated receptor may produce a submaximal signal resulting in reduced colony formation in vitro (23). The numbers of			194,866,973	77%		050%				050%		Results.txt
25	subjects to \approx 1.3% of normal PMN levels. To confirm the EIA results demonstrating reduced 5-LO metabolic capacity in PMN			193,314,899	17%		050%				050%		Results.txt
26	or A \rightarrow V dysfunction (Table 3). These results show that DMPK loss results in A \rightarrow V conduction delay and demonstrate a critical role for			191,377,820	35%		059%				059%		Results.txt
27	had any higher-grade atrial or A \rightarrow V dysfunction (Table 3). These results show that DMPK loss results in A \rightarrow V conduction delay and			191,372,820	13%		059%				059%		Results.txt
28	(specified in the Methods section), and all gave comparable results. Although the immunostaining data confirmed the presence			190,526,783	10%		059%				059%		Results.txt
29	viral replication. Suppression of viral replication by HAART may result in a substantially decreased antigenic stimulus, reduced			193,228,893	24%		050%				050%		Results.txt
30	node, with a lesser decrease in the follicular center areas. These results suggest that viral antigen stimulates immune activation in			193,204,894	16%		050%				050%		Results.txt
31	in DMPK-deficient mice differentially compared with controls. These results indicate that sympatholytic effects are not directly			191,732,833	15%		050%				050%		Results.txt
32	tissue apoptosis and peripheral CD4 T cells (Fig. 3). These results indicate that HAART decreases lymphoid tissue apoptosis			187,262,634	8%		058%				058%		Results.txt

concordancecollocatesplotpatternsclustersfilenamesfollow upsource textnotes

4,006Sete controls, in the absence of added cytokines (Fig. 1, A and C). The number of colonies that grew in the absence of added growth factors is relatively low compared with the number of colonies in the presence of factors. The number of cells pl

Correu eletr3nic : ...

DA\TESIS\CORPUS FL...

Microsoft PowerPoi...

WordSmith Tools

Concord

ES

Corpus Linguistics as a methodological tool for ESP teaching

Most users access a corpus through **concordance lines**, which bring together many instances of a key word in context (**KWIC lines**) and allow the user to observe regularities in use that would remain unseen without a context.

It is through concordance lines that phraseology (i.e., **multiword expressions, clusters, lexical bundles**) is observed.

Corpus Linguistics as a methodological tool for ESP teaching

Phraseology and the use of concordancing programmes are very useful tools **to explain** certain linguistic phenomena, such as **words that are often confused by L2 learners, polysemy, etc.**

Example 4:

Below are a few concordance lines of the adjectives *boring* and *bored* (selected at random from the *BNC*)

Corpus Linguistics as a methodological tool for ESP teaching

Concordance.pdf - Adobe Reader

Archivo Edición Ver Ventana Ayuda

Herramientas Firmar Comentario

Sketch Engine

Search in Help

user: Natalia Laso corpus: British National Corpus

Query (bored)-j 1,323 (11.8 per million)

J2B , so I was encouraged to do P.P.E. I was **bored** by both Philosophy and Economics so I only

J2B secretary or crane driver. When she got **bored** she looked further afield, and took a job

J0R Cynics. It could be ideal, but is it? Dead **bored** , a pink-tongued gorilla picks a meal out

J0W a time that he ad-libbed because he got **bored** with his lines. In the end, you just could

J0W recognize the script. It wasn't that he was **bored** with the show, he was bored with his lines

J0W that he was bored with the show, he was **bored** with his lines. He just got inventive.

J0W between takes, too. `Sometimes, if he got **bored** he would just come up and start reciting

J0W would want to? There are some people who get **bored** who can just fall asleep, standing up with

J0W outrageous at all. I think he had just got **bored** and was looking for something to do, so

J0W dropped because Took and Marty Feldman were **bored** by them. There was such a protest from

J9Y doing nothing, they wouldn't they'd soon get **bored** with it. But when you're self employed

JN7 message. And she's now getting extremely **bored** with that job. It's fairly obviously and

JN7 problem out because I think she's getting **bored** in what she's doing Groups to do what?

JND jump into things first and yeah and get **bored** half way through somebody else Um could

JN8 conference on that Thursday and Friday. He got **bored** silly and considered it a complete waste

JXX was that Ross must have been thoroughly **bored** by her complete lack of sophistication,

Feedback

18:20 05/04/2015

Corpus Linguistics as a methodological tool for ESP teaching

Concordance.pdf - Adobe Reader

Archivo Edición Ver Ventana Ayuda

Herramientas Firmar Comentario

Sketch Engine

Search in Help

user: Natalia Laso corpus: British National Corpus

Query (boring)-j 1,546 (13.8 per million)

J1E the first half: City were playing a very **boring** offside game, Whelan thought he'd got free

J1E Wed, 18 Aug 1993 11:30:13 +0100 Subject: **Boring** Admin Stuff <p> Hi, </p><p> This is a bit boring|

J1E Boring Admin Stuff <p> Hi, </p><p> This is a bit **boring** and non-Leeds, but it may prove useful

J1E 18 Aug 1993 11:49:00 +0100 Subject: RE: **Boring** admin stuff <p> I've heard from Tim and Mark

J1E fans in England' my arse. </p><p> Who was the **boring** twat on radio 5 as guest commentator? I

J1E than West Ham, cos at least Sheff Utd arent **boring** 10 man defence. SCUM are probably kidding

J1E it's now time for me to send one of those **boring** sending off messages that everyone usually

J10 the river at all. He said it was slow and **boring** .' </p><p> `Then why .' </p><p> Nanette shrugged

J10 English tomb-robbers. I was merely making a **boring** journey by river and Luxor happened to

J10 </p><p> `What, then, was the purpose of your **boring** journey?' </p><p> The Prince, unexpectedly

J1C Date: Mon, 17 Jan 94 10:59:40 GMT Subject: **Boring** Saturday <p> I used to look forward to commentary

J1C is I think we'll wind up in one of those **boring** 1-0 or 2-0 defeats where nothing exciting

J1C get slated for not scoring, and for the **boring** game. Also, I might add, compared to scum

J1C match finished with a few variations on ` **Boring** Boring Ipswich' and `I can't read and I

J1C finished with a few variations on `Boring **Boring** Ipswich' and `I can't read and I can't

J1H inspiration and become the predictable and **boring** team to play. Lets hope that we get a decent

Feedback

18:18 05/04/2015

Corpus Linguistics as a methodological tool for ESP teaching

“The importance of corpora lies not in the answers they provide but in the questions they provoke.” (Widdowson 2000:23)