What can I do with Corpus Linguistics?

STELLA E. O. TAGNIN EAGILE - 12/11/2014







This talk

- Introducing Conventionality/Phraseology
 - Different categories
- What is Corpus Linguistics
 Empiricist approach
- How to use Corpus Linguistics
 Corpora online

Convention:

 "a way of behaving that is generally accepted as being normal or right"

• "a way of writing, painting, etc. that is accepted as the usual or right way" (Macmillan English Dictionary)

Convention:

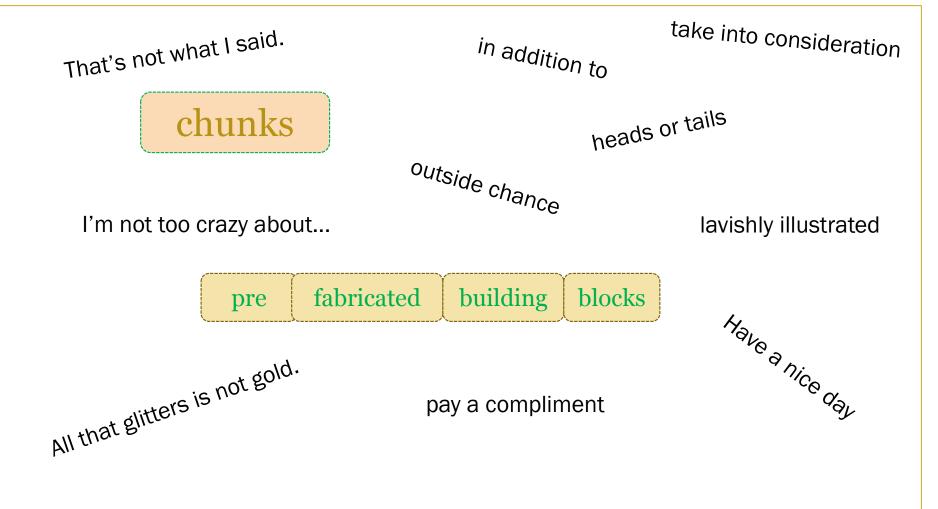
 "a way of behaving that is generally accepted as being normal or right"

• "a way of writing, painting, etc. that is accepted as the **usual** or right way" (Macmillan English Dictionary)

• \rightarrow phraseology = prefabricated chunks







Prepositions: insist on, instructor in, dressed in, good at

Phrasal Verbs: look up, get across, make up

Adjectival Collocations:

public television Japanese current hard luck baking powder

Nominal Collocations:

labor pains seashore cost of living computer science

Verbal Collocations: pay a compliment take a step make a date

Adverbial Collocations:

adverb + adjective lavishly illustrated highly successful

verb + *adverb* lie **outright fully** agree

Binomials and Multinomials :

dead or alive cats and dogs up and down come and go knife and fork

red, white and blue bacon, eggs and coffee men, women and children

Other phraseologies

Idiomatic expressions: kick the bucket; blow one's top

Proverbs: Not all that glitters is gold.

Situational formulas: Act your age, It takes two to tango, Have a nice day.

And even longer ones:

• I, ____, residing at _____, appoint _____, residing at _____ to be my continuing attorney (my "Attorney") in terms of section 15 of the Adults with Incapacity (Scotland) Act 2000 (which act and any subsequent amendment of that is referred to as the "Act"). I declare that I intend that this power of attorney shall be a continuing one which shall continue to have effect even if I shall have lost capacity in relation to the matters in this deed.

This paper aims to explain / provide / present / examine / analyze (11.800.000 hits in Google)

The aim of this article is to explain / analyze / introduce / summarize (49.000.000 hits in Google)



there are many more fixed expressions in language than one would believe!

How can we identify them?

That's where Corpus Linguistics comes into the picture

When incomes are about equal, both are freer to stay or go. [p] 6 Women, more
the world's longest-serving foreign minister, stay or go? What of Mangosuthu
of my life are very important and whether I stay or go remains to be seen [p]
three years are very important and whether I stay or go remains to be seen.
do you think I should do. Do you think I should stay or go? [tc text=laughs]



What is a corpus?

A collection of digitized texts, compiled according to certain criteria, considered to be representative of a language or of the part one intends to study

What is a corpus?

A collection of digitized texts, compiled according to certain criteria, considered to be representative of a language or of the part one intends to study

Corpus linguist vs armchair linguist

"... I don't think there can be any corpora, however large, that contain information about all of the areas of English lexicon and grammar that I want to explore..."

Corpus linguist vs armchair linguist

- "... Every corpus [...], however small, has taught me facts that I couldn't imagine finding out about in any other way."
- (Fillmore, C.J. "Corpus linguistics" or "Computer-aided armchair linguistics". In Svartvik, Jan (ed.) Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991, Berlin/NY: Mouton de Gruyter, 1992, 35-60)

A corpus allows us to

examine a large number of occurrences

of authentic language

at one glance



lexical patterns

grammatical patterns

through CONCORDANCES

Example of concordance: V + or + V

to protest Japan's continued whaling, call **or** write to the Japanese the police know of anything suspicious you see or hear, you are helping to in breast tissue long before a woman can feel **or** notice anything unusual or t even need a stamp. [p] Remember - if you use, or allow other people to use, about AIDS, countering stories that confuse or frighten people or suggest Neither does Midlands Electricity plc confirm or deny any of the claims Uvistat Sun Cream spf15), so the sun won't burn **or** damage her skin. aircraft are a scourge to those who live, work **or** relax in the country. [p] which you lay brick or stone paths can shorten **or** increase their apparent the dark-light hurt my eyes, and I couldn't read **or** watch TV. Others didn't I was pressed by mps of all parties to confirm **or** deny reports that the cost in town. Except for Ella, no one would know **or** care where Sue Anne Norton humankind or the goodwill of God can do to avert **or** transform the effects of was my experience that a patient's will to live (or die) made a substantial be inadequate because it was unable to predict or suggest how managers' credit cards are often useless. You can't bribe or barter with plastic, alcohol or other sedatives. 4. Not drive or operate machinery whilst. should use any opportunity that he has to visit or contact friends or

What's so special about it?

Observation

vs. introspection

Probability

vs. possibility



> Read lines VERTICALLY (top down) to identify patterns

> Read lines HORIZONTALY (across) to identify meanings

Candidates for binomials

to protest Japan's continued whaling, call **or** write to the Japanese the police know of anything suspicious you see or hear, you are helping to in breast tissue long before a woman can feel or notice anything unusual or t even need a stamp. [p] Remember - if you use, or allow other people to use, about AIDS, countering stories that confuse or frighten people or suggest Neither does Midlands Electricity plc confirm or deny any of the claims Uvistat Sun Cream spf15), so the sun won't burn or damage her skin. aircraft are a scourge to those who live, work **or** relax in the country. [p] which you lay brick or stone paths can shorten **or** increase their apparent the dark-light hurt my eyes, and I couldn't read or watch TV. Others didn't I was pressed by mps of all parties to confirm or deny reports that the cost in town. Except for Ella, no one would know or care where Sue Anne Norton humankind or the goodwill of God can do to avert or transform the effects of was my experience that a patient's will to live (or die) made a substantial be inadequate because it was unable to predict or suggest how managers' credit cards are often useless. You can't bribe or barter with plastic, alcohol or other sedatives. 4. Not drive or operate machinery whilst should use any opportunity that he has to visit or contact friends or to a breach of covenant or agreement to keep **or** put in repair any of the written, this novel fails to convince or engage (Abacus £ 6.99). PB two major companies and their decisions can make **or** break small publications.

Let's see what we can do with corpora

THE CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA)

Lácis-Web.





corpus.byu.edu eight corpora | 45-450 million words each

Created by Mark Davies, BYU. Overview, uses, researchers, publications, corpus-based resources.

English	# words	language / dialect	time period	compare to:
Corpus of Contemporary American English (COCA)	450 million	American English	1990-2012	Google, BNC, ANC, BoE COCA/BNC side-by-side
Corpus of Historical American English (COHA)	400 million	American English	1810-2009	Google Books, small corpora
TIME Magazine Corpus of American English	100 million	American English	1923-2006	
Corpus of American Soap Operas	100 million	American English	2001-2012	COCA / BNC
BYU-BNC: British National Corpus*	100 million	British English	1980s-1993	COCA, COCA/BNC side-by-side
Strathy Corpus (Canada)	50 million	Canadian English	~1970s-2000s	
Other languages				
Corpus del Español	100 million	Spanish	1200s-1900s	CORDE and CREA
Corpus do Português	45 million	Portuguese	1300s-1900s	

Online corpora - English

- COCA Corpus of Contemporary American English - <u>www.americancorpus.org/</u>
 - 450 million words
 - 0 1990 2012
 - POS-tagged

BYU-BNC– British National Corpus 100 million words 1980-1993

Online Corpora - Portuguese

Lácio-Web: <u>http://www.nilc.icmc.usp.br/lacioweb</u> 9 million words 8 different domains downloadable

Corpus do Português http://www.corpusdoportugues.org/ 45 million words 1300s – 1900s

Online corpora - bilingual

COMPARA (parallel- English $\leftarrow \rightarrow$ Portuguese)

http://www.linguateca.pt/COMPARA

CoMET: <u>http://www.comet.fflch.usp.br</u> CorTec (comparable – English and Portuguese) CorTrad (parallel – English ← → Portuguese)

What are corpora good for?

Description of the language oLexicon Grammar Semantics Lexicology/Terminology • Dictionaries and glossaries Teaching DDL (data-driven learning) • Error analysis (learner corpus) • Teaching materials

DDL- Data-driven learning

Three steps:

1. Observe



3. Generalize





2. Infer – form hypotheses

confirm or deny

Household had seen them, it was impossible to **confirm** or deny that they of the reason and no one is prepared to **confirm** or deny that the to show on his face that would either **confirm** or deny. When's the merely smiled and nodded and refused to **confirm** or deny anything. Of were, ended there and then. Neither side will **confirm** or deny that they took [p] The Prison Service yesterday refused to **confirm** or deny whether women the Hong Kong police were last night able to confirm or deny the report, by Lord Stevens of Ludgate, has refused to **confirm** or deny claims that it soon. [p] The Agriculture Minister refused to **confirm** or deny reports that he s Playboy magazine Elizabeth Ward refuses to **confirm** or deny whether she has [p] But Downing Street later refused to **confirm** or deny whether was in danger of going blind. [p] Refusing to **confirm** or deny the story, he [sh] Refused [/sh] [p] But Radford refused to **confirm** or deny that Norman had [p] Bosses at smithkline refused refused to **confirm** or deny the sale from her first marriage coyly refused to **confirm** or deny the rumours. 50,000 each. [p] Last night Andrew refused to **confirm** or deny his interest in he was not hurt. The Israeli army would not **confirm** or deny the report. Buckingham Palace is still refusing to **confirm** or deny the split so is

confirm or deny

Household had seen them, it was impossible to **confirm** or deny that they were of the reason and no one is prepared to **confirm** or deny that the to show on his face that would either **confirm** or deny. When's the merely smiled and nodded and refused to **confirm** or deny anything. Of were, ended there and then. Neither side will **confirm** or deny that they took [p] The Prison Service yesterday refused to **confirm** or deny whether women the Hong Kong police were last night able to **confirm** or deny the report, by Lord Stevens of Ludgate, has refused to **confirm** or deny claims that it soon. [p] The Agriculture Minister refused to **confirm** or deny reports that he s Playboy magazine Elizabeth Ward refuses to **confirm** or deny whether she has [p] But Downing Street later refused to **confirm** or denv whether was in danger of going blind. [p] Refusing to **confirm** or deny the story, he [sh] Refused [/sh] [p] But Radford refused to **confirm** or deny that Norman had [p] Bosses at smithkline refused refused to **confirm** or deny the sale from her first marriage coyly refused to **confirm** or deny the rumours. 50,000 each. [p] Last night Andrew refused to **confirm** or deny his interest in he was not hurt. The Israeli army would not **confirm** or deny the report. Buckingham Palace is still refusing to **confirm** or deny the split so is

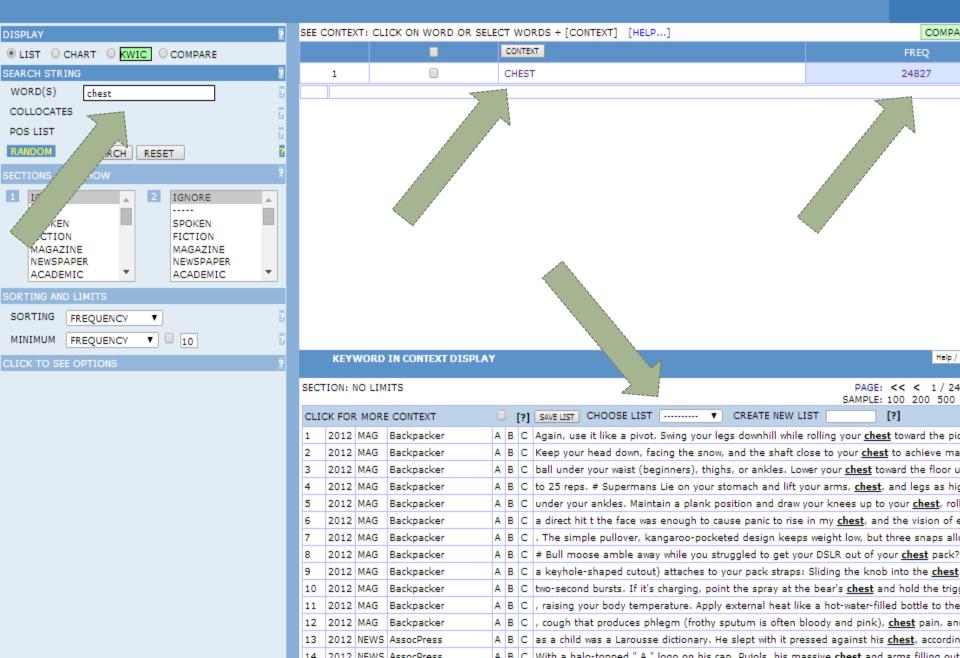
confirm or deny

Household had seen them, it was impossible to confirm or deny that they were of the reason and no one is prepared to **confirm** or deny that the to show on his face that would either **confirm** or deny. When's the merely smiled and nodded and refused to **confirm** or deny anything. Of were, ended there and then. Neither side will **confirm** or deny that they took [p] The Prison Service yesterday refused to **confirm** or deny whether women the Hong Kong police were last night able to **confirm** or deny the report, by Lord Stevens of Ludgate, has refused to **confirm** or deny claims that it soon. [p] The Agriculture Minister refused to **confirm** or deny reports that he s Playboy magazine Elizabeth Ward refuses to **confirm** or deny whether she has [p] But Downing Street later refused to **confirm** or deny whether was in danger of going blind. [p] Refusing to **confirm** or deny the story, he [sh] Refused [/sh] [p] But Radford refused to **confirm** or deny that Norman had [p] Bosses at smithkline refused refused to **confirm** or deny the sale from her first marriage coyly refused to confirm or deny the rumours. 50,000 each. [p] Last night Andrew refused to **confirm** or deny his interest in he was not hurt. The Israeli army would not **confirm** or deny the report. Buckingham Palace is still refusing to **confirm** or deny the split so is

CORPUS OF CONTEMPORARY AMERICAN ENGLISH

450 MILLION WORDS, 1990-2012 [DOWNLOAD ALL 190,000 TEXTS]

-- START -- Y ?



Concordance for *chest*

to 25 reps. # Supermans Lie on your stomach and lift your arms, **<u>chest</u>**, and legs as high as possible. Hold for several seconds. (Alternative

Bull moose amble away while you struggled to get your DSLR out of your <u>chest</u> pack? Nab the next shot with this nifty system. " It's the

a keyhole-shaped cutout) attaches to your pack straps: Sliding the knob into the **<u>chest</u>** plate stabilizes the camera when worn with a neck strap. Keep a rain cover

two-second bursts. If it's charging, point the spray at the bear's <u>chest</u> and hold the trigger until the can is fully discharged. Out of spray and

. " # A police report shows the 17-year-old had been shot once in the **<u>chest</u>** and had been pronounced dead at the scene. The autopsy says the fatal shot

, rich and poor, the gentrifiers and the gentrified. Mr. Jeffries's campaign <u>chest</u> is many times the size of Mr. Barron's, having collected by this week

and the \$8 million, privately funded campaign for it might be the biggest war **<u>chest</u>** in the history of such measures. # And opponents mean to stop it.

22, a Park Police officer was kicked in the groin and another in the **<u>chest</u>** when they attempted to make an arrest in McPherson Square, according to court documents

up to east Austin, and that Doggett - with his \$3 million campaign war **<u>chest</u>** - was now her opponent. # " He made the decision to challenge me

Concordance for *chest*

to 25 reps. # Supermans Lie on your stomach and lift your arms, **<u>chest</u>**, and legs as high as possible. Hold for several seconds. (Alternative

Bull moose amble away while you struggled to get your DSLR out of your <u>chest</u> pack? Nab the next shot with this nifty system. " It's the

a keyhole-shaped cutout) attaches to your pack straps: Sliding the knob into the chest plate stabilizes the camera when worn with a neck strap. Keep a rain cover

two-second bursts. If it's charging, point the spray at the bear's <u>chest</u> and hold the trigger until the can is fully discharged. Out of spray and

. " # A police report shows the 17-year-old had been shot once in the **<u>chest</u>** and had been pronounced dead at the scene. The autopsy says the fatal shot

, rich and poor, the gentrifiers and the gentrified. Mr. Jeffries's campaign <u>chest</u> is many times the size of Mr. Barron's, having collected by this week

and the \$8 million, privately funded campaign for it might be the biggest war <u>chest</u> in the history of such measures. # And opponents mean to stop it.

22, a Park Police officer was kicked in the groin and another in the **<u>chest</u>** when they attempted to make an arrest in McPherson Square, according to court documents

up to east Austin, and that Doggett - with his \$3 million campaign war <u>chest</u> - was now her opponent. # " He made the decision to challenge me

CORPUS OF CONTEMPORARY AMERICAN ENGLISH

450 MILLION WORDS, 1990-2012 [DOWNLOAD ALL 190,000 TEXTS]

-- START -- Y ?

A B C comments suggest they aren't yet marching behind the candidate with the biggest war ch

DISPLAY	SEE CONTEXT: CLICK	ON WORD OR SE	LECT WORDS + [CONTEXT] [HELP]	COMPARE	
● LIST ○ CHART ○ KWIC ○ COMPARE		•	CONTEXT	FREQ	
SEARCH STRING	1		PAIN	479	
WORD(S) chest	2		PAINS	244	
COLLOCATES [nn*] 1 V 1 V	3		WAR	215	
POS LIST noun.ALL	4		X-RAY	212	
RANDOM SEARCH RESET	5		ICE	180	
SECTIONS SHOW ?	6		TREASURE	151	
1 IGNORE SPOKEN FICTION MAGAZINE NEWSPAPER ACADEMIC	7		MEDICINE	124	
	8		HAIR	118	
	9		WALL	106	
	10	0	CAVITY	96	
	11	0	MUSCLES	83	
SORTING AND LIMITS	12	-	HEAVING	75	
SORTING FREQUENCY V	13		BARREL	69	
MINIMUM FREQUENCY V 10	14		X-RAYS	67	
CLICK TO SEE OPTIONS		CONTEXT DISPLA		Help / in	
	SECTION: NO LIMITS PAGE: << < 1 / 3 > SAMPLE: 100 200				
	CLICK FOR MORE C	ONTEXT	[?] SAVE LIST CHOOSE LIST	[?]	
	1 2012 SPOK Fo	x_Hannity /	A B C if we have enough time to raise money to go up against his billion war g	<u>:hest</u> that he inten	
	2 2012 SPOK NP	R_ATC	A B C of independent campaign organization. And he was looking to basically	build a campaign <u>w</u>	
	3 2012 NEWS At		A B C and the \$8 million, privately funded campaign for it might be the bigges		
	4 2012 NEWS Wa		 A B C up to east Austin, and that Doggett - with his \$3 million campaign war c A B C hand, Facebook's IPO was a financial success. The company raised a war 		
	6 2011 SPOK Fo		A B C nand, Facebook's IPO was a financial success. The company raised a <u>wa</u> A B C people in at their own cost. I mean, Romney has the biggest war chest .		
	7 2011 SPOK Fo	-	B C in front of the GOP pack with more than \$18 million in his campaign war		
	8 2011 SPOK CE	- BS_Early	A B C field among Republican voters. Perry has more than fifteen million dolla		
	9 2011 SPOK NB	C_Matthews	A B C front-runner. He's going to have a lot of money in his campaign war che	<u>est</u> , and he got out	
	10 2011 SPOK PB	-	A B C to Florida tonight to fund-raise, because he needs to build up a huge wa		
			A B C through the park, dumping everything on the crab grass. Inside was a \underline{w}		
	12 2011 MAG Co		A B C 888525118 With a <u>war chest</u> approaching \$33 million, Britain's Ministry		
	13 2011 NEWS CS	Monitor /	A B C $$ level the playing field " between candidates able to assemble a sizable	e campaign <u>war</u> <u>ch</u> e	

14 2011 NEWS SanEranChron

Concordance for *war chest*

and the \$8 million, privately funded campaign for it might be the biggest **war <u>chest</u>** in the history of such measures. # And opponents mean to stop it.

hand, Facebook's IPO was a financial success. The company raised a <u>war</u> <u>chest</u> of \$16 billion and minted a small army of millionaires and billionaires, capping off

field among Republican voters. Perry has more than fifteen million dollars in his **war chest**, giving him plenty of ammunition to turn things around and be the viable option

front-runner. He's going to have a lot of money in his campaign **war chest**, and he got out of this debate unscathed. I still think that his

through the park, dumping everything on the crab grass. Inside was a <u>war</u> <u>chest</u>. During summers, the trees were like tinder, with calls in outlying areas

" level the playing field " between candidates able to assemble a sizable campaign <u>war</u> <u>chest</u> and those who are unable or unwilling to engage in private campaign fundraising. #

Cain has enough time and resources to build a full-fledged organization. # His **war chest** is dwarfed by those of his opponents. Former Massachusetts Gov. Mitt Romney has

in the head, a popular two-term senator with plenty of money in his **war chest** and ahead in most of the polls taking himself out of the running for reelection

-- if they simply eliminated the dividend, they would have a \$10 billion-a-year war chest to clean things up. And then, also, BP has another \$25 billion

. On May 10, Europeans bailed out Greece, authorized a euro750 billion <u>war</u> <u>chest</u> to protect currencies, and authorized the European Central Bank to intervene to buy up

All of this with a view to squelching dissent and plumping up Chavez's **war chest** for the 2012 elections, when he will stand for reelection. For a nation

Concordance for *war chest*

and the \$8 million, privately funded campaign for it might be the biggest **war chest** in the history of such measures. # And opponents mean to stop it.

hand, Facebook's IPO was a financial success. The company raised a <u>war</u> <u>chest</u> of \$16 billion and minted a small army of millionaires and billionaires, capping off

field among Republican voters. Perry has more than fifteen million dollars in his war chest, giving him plenty of ammunition to turn things around and be the viable option

front-runner. He's going to have a lot of money in his campaign war chest, and he got out of this debate unscathed. I still think that his

through the park, dumping everything on the crab grass. Inside was a war chest. During summers, the trees were like tinder, with calls in outlying areas

" level the playing field " between candidates able to assemble a sizable campaign war chest and those who are unable or unwilling to engage in private campaign fundraising. #

Cain has enough time and resources to build a full-fledged organization. # His <u>war</u> <u>chest</u> is dwarfed by those of his opponents. Former Massachusetts Gov. <u>Mitt Romney</u> has

in the head, a popular two-term senator with plenty of money in his **war chest** and ahead in most of the polls taking himself out of the running for reelection

-- if they simply eliminated the dividend, they would have a \$10 billion-a-year war chest to clean things up. And then, also, BP has another \$25 billion

. On May 10, Europeans bailed out Greece, authorized a euro750 billion <u>war</u> chest to protect currencies, and authorized the European Central Bank to intervene to buy up

All of this with a view to squelching dissent and plumping up Chavez's war chest for the 2012 elections, when he will stand for reelection. For a nation



A **war chest** is a collection of funds intended to allow a person or organization to get through a situation that requires much more money than usual.

> In arms and armor, a **war chest** is a container for the personal weapons and protective gear of a citizensoldier, kept in the household, and is the origin of the term. The term's modern meaning originates with the medieval practice of having a chest, literally, filled with money to open in time of war.

CORPUS OF CONTEMPORARY AMERICAN ENGLISH

450 MILLION WORDS, 1990-2012 [DOWNLOAD ALL 190,000 TEXTS]

DISPLAY	SEE CONTEXT: CLICK C	N WORD OR SEL	ECT WORDS + [CONTEXT] [HELP]	COMPARE 7
● LIST ○ CHART ○ KWIC ○ COMPARE			CONTEXT		FREQ
SEARCH STRING	1		HAVE		4004
WORD(S) opportunit*	2		HAD		2606
COLLOCATES [v*] 2 V 0 V			PROVIDE		1213
POS LIST verb.ALL V	4		IS		1188
KANDOM SEARCH RESET	5		GIVEN		1124
SECTIONS SHOW ?	6		'S		609
1 IGNORE	7		PROVIDES		597
SPOKEN SPOKEN	8		HAS		573
FICTION FICTION MAGAZINE MAGAZINE	9		WAS		527
NEWSPAPER NEWSPAPER	10		ARE		525
	11		ТАКЕ		501
SORTING AND LIMITS	12		BE		492
SORTING FREQUENCY V	13		PROVIDING		436
MINIMUM FREQUENCY V 10	14		GET		425
CLICK TO SEE OPTIONS	KEYWORD IN C	ONTEXT DISPLA	ř		Help / information ,
	SECTION: NO LIMITS			7	PAGE: << < 1 / 13 > >> SAMPLE: 100 200 500 1000
	CLICK FOR MORE CON	ITEXT	[?] SAVE LIST CHOO	OSE LIST V CREATE NEW LIST	[?]
	1 2012 SPOK NBC	_MeetPress	A B C your judgment?	MR-AXELROD: This president guarantees an e	economy that will grow and provide
	2 2012 SPOK PBS	-		w, I'm sure there are places that will provide th	
	3 2012 NEWS Den			e letting it off the lot. Plus, those short classes	
		-		n presentations, students explain their maps.	
	5 2012 ACAD Prac			dvance of travel with specific food requests and quently (e.g., only once a month) may not pro	
		2 1		quently (e.g., only once a month) may not pro ffective teaching techniques and research-base	
	8 2012 ACAD Teac			nd Creativity (http: **44;129352;TOOLONG). 1	
				at represented by direct transmissions of STLV	
	S EVIE NOND LINE	gingrineerous	n e e ee come the	serve and a server of an external manons of or ever	a because may can provide an or

11 2012 ACAD PracticeNurse

14 2011 MAG TodaysParent

12 2011 SPOK NPR_TalkNation

10 2012 ACAD AmericanSecondary A B C complement the skills taught at school. Home settings can, however, provide realistic opportunit

13 2011 FIC Bk:DevilInDisguise A B C 's false assumption. On the contrary, she'd worked feverishly to provide every opportunity for Lu

A B C point in the future. Annual review appointments or flu immunisation clinics may provide suitable

A B C That was a great deal of the homeownership push was to try to provide homeownership opportun

A B C situation, which reinforces her fear of interacting with others. Try to provide gentle opportunities

-- START -- 🔻 ?

1	HAVE	4004	
2	HAD	2606	have + had + has + having = 7.379
3	PROVIDE	1213	
4	IS	1188	
5	GIVEN	1124	provide + provides + providing + provided = 2.62
6	'S	609	
7	PROVIDES	597	
8	HAS	573	
9	WAS	527	
10	ARE	525	
11	ТАКЕ	501	
12	BE	492	
13	PROVIDING	436	
14	GET	425	
15	CREATE	382	
16	PROVIDED	381	
17	OFFER	380	
18	OFFERS	342	
19	MISSED	302	
20	ТООК	301	
21	SAW	287	
22	SEIZE	275	
23	SEE	249	
24	OFFERED	246	
25	CREATING	208	
26	LEARNING	198	
27	HAVING	196	
28	MISS	189	
29	SEIZED	183	
30	GOT	163	

1	HAVE	4004
2	HAD	2606
3	PROVIDE	1213
4	IS	1188
5	GIVEN	1124
6	'S	609
7	PROVIDES	597
8	HAS	573
9	WAS	527
10	ARE	525
11	TAKE	501
12	BE	492
13	PROVIDING	436
14	GET	425
15	CREATE	382
16	PROVIDED	381
17	OFFER	380
18	OFFERS	342
19	MISSED	302
20	ТООК	301
21	SAW	287
22	SEIZE	275
23	SEE	249
24	OFFERED	246
25	CREATING	208
26	LEARNING	198
27	HAVING	196
28	MISS	189
29	SEIZED	183
30	GOT	163

Can someone *give* you an opportunity???

're celebrating four years of Title IX and it shows when girls are **given** the **opportunity** and they're competing and they're working as a team, it makes them

that we love most. And so I'm so grateful to be **given** another **opportunity** to sing and to share. And it brings me more joy than I can

Commandments - but those people who tried to recall the Ten Commandments, **given** the **opportunity** to cheat, did not cheat at all. It wasn't that the more

Winters might drink beer from a bottle, I would probably never be **given** the **opportunity** to sleep with her again. She's kinda weird that way. # "

tail against the floor. He'd been known to steal things, **given** the **opportunity**, and had recently taken a liking to Hook's prosthesis. Just last week

saying nothing. # " Anyway, it seems that you are being **given** an **opportunity** to help America and prove that you are indeed loyal. The men from Washington

Black women show that they have what it takes, they are being **given** more **opportunities** to climb. # Q What advice do you have for Black women ascending to

HAVE	4004
HAD	2606
PROVIDE	1213
IS	1188
GIVEN	1124
'S	609
PROVIDES	597
HAS	573
WAS	527
ARE	525
ТАКЕ	501
BE	492
PROVIDING	436
GET	425
CREATE	382
PROVIDED	381
OFFER	380
OFFERS	342
MISSED	302
ТООК	301
SAW	287
SEIZE	275
SFF	249
OFFERED	246
CREATING	208
LEARNING	198
HAVING	196
MISS	189
SEIZED	183
GOT	163
	HADPROVIDEISGIVEN'SHASHASHASARETAKEBEPROVIDINGBEPROVIDINGOFFEROFFEROFFERSOFFERSOFFERSSEIZESEIZESEIZESEINISSEDLEARNINGHAVINGMISSMISSSEIZE



PROJETO CO rpus M ultilíngue para E nsino e

Tradução

Projeto CoMET - FFLCH/USP

Av Prof. Luciano Gualberto, 403. 3º andar - sala 14 (mapa) Cidade Universitária - Butantã 05508-010 São Paulo, SP - Brasil



			CORTEC
O que é	Lista de corpora	Pesquisar	

Av Prof. Luciano Gualberto, 403. 3º andar - sala 14 (mapa) Cidade Universitária - Butantã 05508-010 São Paulo, SP - Brasil 11 3091-2929 projetocomet @ usp.br



Search What It Is

Step 1/3: Corpus selection.

Select which one(s) you want to work on.

List of Corpora

Culinária

- Ecoturismo
- 🗆 Hipertensão
- Informática
- □ Instrumentos Contratuais
- Astronomia
- Insuficiência Renal
- 🗆 Linguística
- Medidores Eletromagnéticos de Vazão
- Suplementos Nutricionais
- Computação
- Café
- Turismo Cultural
- 🗆 Culinária 2
- Prostodontia
- 🗆 Fotografia
- Autoclaves
- 🗆 Moda
- Turismo Hotelaria
- Futebol

Language

O Portuguese

English

*

Português



	Po	rtuguês
Search	What It Is	
Ste	p 2/3: Select your tool.	
L	st of Tools	
(Concordancer Frequency Counters N-Gram Generator	
	>> Next Step	

CorTec – Technical Scientific Corpus

Login



Português

Search What It Is

Step 3/3: Applying Tool to Corpus.

Select the tool according to the options below:

Concordancer	
Expression or word: Same as v pepper	
Same as	
Case sensitive: Starting with	
Ending in	
Size of reduced context (in characters): 30 •	
	>> Go

Select another tool <<

CorTec – Technical Scientific Corpus



Search What Is CorTec

Foram encontradas 1128 ocorrências!! DOWNLOAD hits.

Click on search word to obtain expanded context (150 characters).

1 2 oz) Cinnamon - 1tsp Black pepper - 1/4 tsp Nutmeg - 1/4 tsp 2 stard powder - 1 tsp Cayenne pepper + pinch Butter - 50g (2 oz) 3 ur, salt, mustard and cayenne pepper into a bowl. Rub the butter i 4 stard powder - 1 tsp Cayenne pepper + pinch Butter - 50g (2 oz) 5 ur, salt, mustard and cayenne pepper into a bowl. Rub butter into 6 tsp Salt - 1/4 tsp Cayenne pepper + shake Butter - 65g (21/2 o 7 tard powder, salt and cayenne pepper into a bowl. Cut the butter i 8 elery salt - 1/2 tsp Cayenne pepper + 1/2 tsp Milk - 1 tbsp MET 9 oz) Salt and freshly ground pepper Oregano - 1 tsp Mushrooms -10 - 225g (8 oz) Salt - 1 tsp Pepper Fresh sage - 2 tbsp, chopped 11 ed. 3. Put the flours, salt, pepper and sage in a large bowl and Westmorland Pepper Cake If you have never tried 12 ool. 13 f you have never tried adding pepper to a sweet dish before now, y 14 ge pinch Finely ground black pepper - 1/2 tsp Milk - 4 tbsp Egg 15 ly. Put the flour, spices and pepper in a bowl and gently stir in 16) Salt Freshly ground black pepper METHOD 1. Put the fish in a 17 ce, season well with salt and pepper then cover with a large shee 18 arge Salt and freshly ground pepper, Sunflower oil - 1 tbsp Onio 19 inside and out with salt and pepper. 2. Heat the oil in a fryi 20 Salt and fresh ground black pepper Parsley sprig - to garnish 21 k Salt Freshly ground black pepper Lemon juice - 4 tsp Cucumbe 22 Salt and freshly ground black pepper Lemon juice - 1 tsp Parsley 23 sprinkle with flour, salt and pepper Lay the fish in the centre. 24 ide of the fish with salt and pepper sprinkle with lemon juice an



Início

CorTrad

	Início	Jornalístico	Literário	▼ Técnico-científico
Bem-vindo ao CorTrad				
O CorTrad é o corpus paralelo de tradução (português-inglês) de paralelos, o CorTrad dispõe de pelo menos duas funcionalidades (original, versões revisadas e tradução publicada); (ii) mecanismo seções específicas dos diferentes tipos textuais.	; inovadoras: (i) a possil	oilidade de se comparar	diferentes versões de u	ım mesmo texto

O CorTrad é um corpus aberto e conta atualmente com três subcorpora:

- CorTrad jornalístico (por ora, divulgação científica)
- · CorTrad literário (por ora, contos)
- CorTrad técnico-científico (por ora, culinária)

A lista das obras incluídas, assim como os agradecimentos devidos, encontram-se em Agradecimentos.

A disponibilização do CorTrad na rede é um projeto conjunto COMET/NILC/Linguateca, usando o sistema DISPARA.

Comentários ou questões para a equipe do CorTrad

Q



Início

CorTrad

•	•	•	•
Início	Jornalístico	Literário	Técnico-científico

CorTrad literário contos

O CorTrad é um corpus aberto, sujeito a alterações. Veja dados quantitativos para informações atualizadas sobre o conteúdo do corpus.

A parte literária do CorTrad consiste atualmente de contos em inglês traduzidos para o português por alunos do Curso de Especialização em Tradução lato sensu da Universidade de São Paulo. A tradução dos 28 contos australianos foi proposta como trabalho final da disciplina Tradução Literária, ministrada por Stella E. O. Tagnin, no 2º semestre de 2003 e publicada em 2005. A tradução dos 20 contos canadenses foi realizada em 1991 e 1992, e publicada em 2002.

A disponibilização do CorTrad na rede é um projeto conjunto entre o COMET, a Linguateca e o NILC, usando o sistema DISPARA.

Pesquisar no corpus

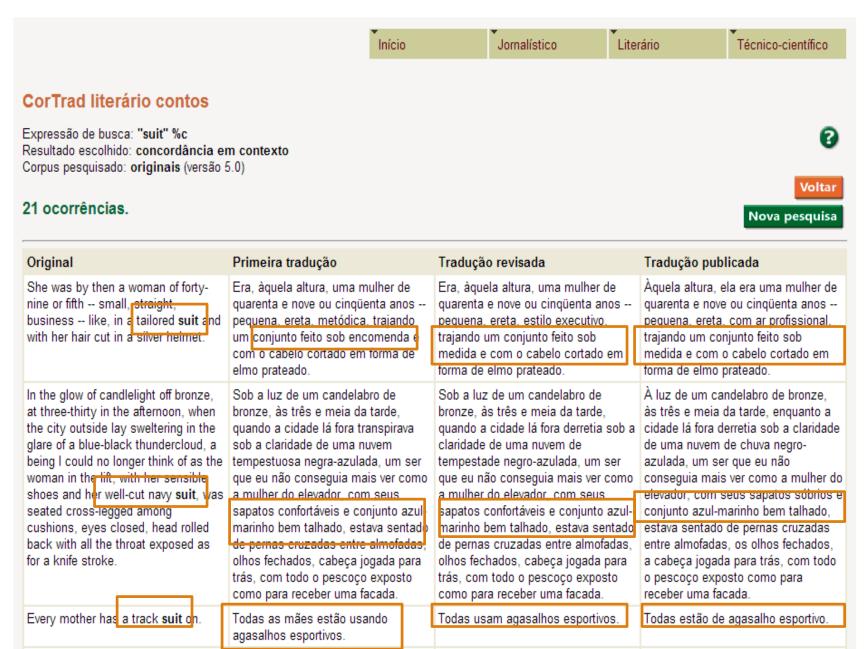
R

Original	o principal suit	✓ ver
Primeira tradução	O principal	✓ ver
Tradução revisada	O principal	✓ ver
Tradução publicada	principal	✓ ver

Ignorar maiúsculas/minúsculas

	Pesquisar
Resultado	
 Concordância 	 Distribuição das formas
 Distribuição dos lemas 	 Distribuição da categoria gramatical (PoS)
 Distribuição do tempo verbal e/ou do caso pronominal 	 Distribuição de pessoa e/ou número
 Distribuição do gênero morfológico 	 Distribuição da função sintática

CorTrad



r rojeto Equipe	
-----------------	--

oonice oonnaa

CorTrad

Início	Jornalístico	Literário	Técnico-científico
--------	--------------	-----------	--------------------

CorTrad literário contos

O CorTrad é um corpus aberto, sujeito a alterações. Veja dados quantitativos para informações atualizadas sobre o conteúdo do corpus.

A parte literária do CorTrad consiste atualmente de contos em inglês traduzidos para o português por alunos do Curso de Especialização em Tradução lato sensu da Universidade de São Paulo. A tradução dos 28 contos australianos foi proposta como trabalho final da disciplina Tradução Literária, ministrada por Stella E. O. Tagnin, no 2º semestre de 2003 e publicada em 2005. A tradução dos 20 contos canadenses foi realizada em 1991 e 1992, e publicada em 2002.

A disponibilização do CorTrad na rede é um projeto conjunto entre o COMET, a Linguateca e o NILC, usando o sistema DISPARA.

ß

Original	O principal	✓ ver
Primeira tradução	O principal	✓ ver
Tradução revisada	O principal	✓ ver
Tradução publicada	or principal conjunto	✓ ver

Ignorar maiúsculas/minúsculas

Pesquisar

Resultado	
Concordância	 Distribuição das formas
 Distribuição dos lemas 	 Distribuição da categoria gramatical (PoS)
 Distribuição do tempo verbal e/ou do caso pronominal 	 Distribuição de pessoa e/ou número
 Distribuição do gênero morfológico 	 Distribuição da função sintática
 Distribuição por obra 	 Distribuição por autor

CorTrad literário contos

Expressão de busca: "conjunto" %c Resultado escolhido: concordância em contexto Corpus pesquisado: tradução final (versão 5.0)

5 ocorrências.

What a way to louse up this new

shiny shnozzola!

magenta outfit - streaming eyes, a

0

Nova pesquisa

Isso é jeito de estragar esse novo

narigão brilhante!

conjunto magenta: olhos vermelhos,

			nova pesquisa
Original	Primeira tradução	Tradução revisada	Tradução publicada
She was by then a woman of forty- nine or fifth small, straight, business like, in a tailored suit and with her hair cut in a silver helmet.	Era, àquela altura, uma mulher de quarenta e nove ou cinqüenta anos pequena, ereta, metódica, trajando um conjunto feito sob encomenda e com o cabelo cortado em forma de elmo prateado.	Era, àquela altura, uma mulher de quarenta e nove ou cinqüenta anos pequena, ereta, estilo executivo, trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.	Àquela altura, ela era uma mulher de quarenta e nove ou cinqüenta anos pequena, ereta, com ar profissional, trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.
In the glow of candlelight off bronze, at three-thirty in the afternoon, when the city outside lay sweltering in the glare of a blue-black thundercloud, a being I could no longer think of as the woman in the lift, with her sensible shoes and her well-cut navy suit, was seated cross-legged among cushions, eyes closed, head rolled back with all the throat exposed as for a knife stroke.	Sob a luz de um candelabro de bronze, às três e meia da tarde, quando a cidade lá fora transpirava sob a claridade de uma nuvem tempestuosa negra-azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos confortáveis e conjunto azul- marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, olhos fechados, cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.	Sob a luz de um candelabro de bronze, às três e meia da tarde, quando a cidade lá fora derretia sob a claridade de uma nuvem de tempestade negro-azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos confortáveis e conjunto azul- marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, olhos fechados, cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.	À luz de um candelabro de bronze, às três e meia da tarde, enquanto a cidade lá fora derretia sob a claridade de uma nuvem de chuva negro- azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos sóbrios e conjunto azul-marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, os olhos fechados, a cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.
There is no story, no set of events that leads anywhere or proves anything no middle, no end.	Não há história, nem conjunto de acontecimentos que leve a algum lugar ou prove alguma coisa nenhum meio, nenhum fim.	Não há história, nem conjunto de acontecimentos que leve a algum lugar ou prove alguma coisa nenhum meio, nenhum fim.	Não há história nem um conjunto de acontecimentos que leve a algum lugar ou prove alguma coisa nenhum meio, nenhum fim.
Altogether Aunt Thalia was like a masked figure in a tragedy.	No conjunto, Tia Thalia parecia uma personagem mascarada de tragédia.	No conjunto, Tia Thalia parecia uma personagem mascarada numa tragédia.	No conjunto , Tia Thalia parecia uma personagem mascarada numa tragédia.

Isso é jeito de estragar esse novo

narigão brilhante!

conjunto magenta: olhos vermelhos,

Isso é jeito de estragar esse novo

narigão brilhante!

conjunto magenta: olhos vermelhos,

Voltar

CL fosters learner autonomy

>You

- > have resource to online corpora
- can carry out your own searches clear doubts
- can observe data
- > can make inferences raise hypotheses
- > can come to your own conclusions

Teacher becomes a mediator

CL in teaching

> Give your students concordance lines to study

- First of all: Show them how to read a concordance
- > Then, ask questions to direct observation
- > In the case of **chest**, for example:
 - Does the word *chest* have the same meaning in all examples?
 - > How many different meanings can you identify? Which ones?
 - > Which, do you believe, is the most common meaning?

Exercises with concordances

Replace the search word with an invented word, like troove, screep.

vely repels any attention that will not screep to its level. I suspect this nistration has found that China's rapid screep in global standing has created tem. Alternatively, there may be a slow screep in the concentrations of these avid Yosifon and Peter N. Steams, " The Screep and Fall of American Posture, " ch policies were closely related to the screep of the welfare state, which came ed by neoliberal economic thinking, the screep of cheap foreign labor, growing dic surgeons saw a new condition on the screep; adolescent idiopathic scoliosis. ermine your roof pitch by measuring the screep over the run using Table 2. For ld of warcraft phishing attempts on the screep. Wired (apr. 29, 2009); http: se New cases of tuberculosis are on the screep in fuel costs has spured eople aged 85 and over. # HPV linked to screep in oral cancer cases The number of

Blank out the search word
 Mix similar words and ask students to identify which is which

What's the difference between *rise* and *raise*?

Concordance for *raise*

600 students, will attract tourism and raise local living standards. To raise m oard's comments have " done nothing but raise additional questions. " # Dambly i rty leaders in Washington that he could raise enough money to become a viable ca ral regulations so they can more easily raise capital, hit a wall after Senate R and other media attention, have helped raise awareness about treatment and led ney is used to cover what Wales can not raise itself from taxes and borrowing. # ar in Afghanistan and a 1.7 percent pay raise for military personnel. The House IG Metall, won a deal for a 4.3 percent raise over 13 months in a key region in een memory and the disk device and then raise an interrupt when it is done. For yed " absolutely no role' in efforts to raise money from the rabbi's followers. f the rabbi's top aides, Ofer Biton, to raise campaign money from the rabbi's fo largely from C 84644 When is it fair to raise the bar for a building? To expect e work with ACM Europe and ACM India to raise awareness of the professional oppo sm and raise local living standards. To raise money, the Rinpoche traveled acros e Future " Mags, originally released to raise money for the Michael J. Fox Found , " a practice of using the Internet to raise capital from a large number of sma ment might affect the club's efforts to raise additional cash. Earlier, he said while Santorum in particular sought to raise his own profile. # Introduced to t n the Arabian Peninsula will be able to raise funds, win recruits and increase t " But it would be more irresponsible to raise the debt ceiling without taking dr solid conservative credentials works to raise his profile beyond his home state Romney will again confront the need to raise the country's borrowing limit or e solid conservative credentials works to raise his profile beyond Florida, if not ies in the Midwest. # "What it did was raise the level of awareness, " said Nor to wither on the vine because it would raise premiums, forcing many seniors to

Concordance for *rise*

diabetics began to rise considerably, a rise which shows no sign of abating. It diabetics began to rise considerably, a rise which shows no sign of abating. It Nocturnal polyuria: there is normally a rise in the secretion of arginine vasopr ines in forest cutting coincided with a rise in imported timber. # In 2010 the W Internet has already been blamed for a rise in divorce as people search social e consequence of this divide could be a rise in rates of divorce and separation, stem has encouraged the development and rise of other powers, from Europe to Jap ion through this past January. Brazil's rise has solidified its reputation as a ticated conspiracy to frustrate China's rise. Washington, meanwhile, has increas ating perceptions of Beijing's economic rise and Washington's relative decline, random variation rather than a genuine rise or decline in citations received by e aspects of American culture that give rise to behavioral risk factors associat e pitch. If a roof measures a four inch rise at a point twelve inches in then th y blurred, crime levels will inevitably rise. And as people increasingly treat e vironmental conditions (e.g., sea-level rise, land degradation and loss, declini ng water shortages, flooding, sea-level rise, and extreme weather events. For ex ecent estimates indicate that sea-level rise of? 1 m could occur by the end of t vely repels any attention that will not rise to its level. I suspect this account nistration has found that China's rapid rise in global standing has created enha tem. Alternatively, there may be a slow rise in the concentrations of these mark avid Yosifon and Peter N. Steams, " The Rise and Fall of American Posture, " The ch policies were closely related to the rise of the welfare state, which came in ed by neoliberal economic thinking, the rise of cheap foreign labor, growing ine dic surgeons saw a new condition on the rise; adolescent idiopathic scoliosis. # ermine your roof pitch by measuring the rise over the run using Table 2. For ins ld of warcraft phishing attempts on the rise. Wired (apr. 29, 2009); http: **86; se New cases of tuberculosis are on the rise according to new figures published 1 (American Airlines Cargo, 2011). This rise in fuel costs has spurred initiativ eople aged 85 and over. # HPV linked to rise in oral cancer cases The number of

Which block is *rise* and which one is *raise*?

diabetics began to rise considerably, a stem has encouraged the development and ion through this past January. Brazil's random variation rather than a genuine e aspects of American culture that give y blurred, crime levels will inevitably vironmental conditions (e.g., sea-level vely repels any attention that will not avid Yosifon and Peter N. Steams, " The to cyclones and storms are expected to

1

2 3

4

5

6

7

8

9

10

1

2

3

4

5

6

7

8

9

10

600 students, will attract tourism and oard's comments have " done nothing but rty leaders in Washington that he could ral regulations so they can more easily and other media attention, have helped ar in Afghanistan and a 1.7 percent pay IG Metall, won a deal for a 4.3 percent e work with ACM Europe and ACM India to sm and raise local living standards. To n the Arabian Peninsula will be able to which shows no sign of abating. It of other powers, from Europe to Jap has solidified its reputation as a or decline in citations received by to behavioral risk factors associat . And as people increasingly treat e , land degradation and loss, declini to its level. I suspect this accoun and Fall of American Posture, " The because of increasing urban poor po

local living standards. To raise m additional questions. " # Dambly i enough money to become a viable ca capital, hit a wall after Senate R awareness about treatment and led for military personnel. The House over 13 months in a key region in awareness of the professional oppo money, the Rinpoche traveled acros funds, win recruits and increase t

Rise or raise???

ar in Afghanistan and a 1.7 percent pay ticated conspiracy to frustrate China's yed " absolutely no role' in efforts to umbers of obese people is continuing to se New cases of tuberculosis are on the , " a practice of using the Internet to while Santorum in particular sought to ies in the Midwest. # " What it did was n rank are indicative of journals' true tem. Alternatively, there may be a slow for military personnel. The House .Washington, meanwhile, has increas money from the rabbi's followers. steadily, with obvious concern for according to new figures published capital from a large number of sma his own profile. # Introduced to t the level of awareness, " said Nor or fall relative to others or the r in the concentrations of these mark

Final remarks

Corpus Linguistics

- will raise your linguistic awareness
- will help you clear your doubts
- will help you develop your research skills
- will foster your learning autonomy
- will provide input for building teaching materials
- will hopefully motivate you and your (future) students to learn "natural and fluent" English

Thank you very much

Stella