



# Using Corpus Linguistics Tools in the ESL Writing Classroom

**Larissa Goulart**  
*Montclair State University*





# LITERACIES FOR ALL:

Where We've Been & Where We're Going



# LITERACIES FOR ALL:

Where We've Been & Where We're Going

## **Academic Literacies**



# LITERACIES FOR ALL:

Where We've Been & Where We're Going

## Academic Literacies

Learning in higher education involves adapting to new ways of knowing: new ways of understanding, interpreting and organising knowledge. Academic literacy practices—reading and writing within disciplines—constitute central processes through which students learn new subjects and develop their knowledge about new areas of study. A practices approach to literacy takes account of the cultural and contextual component of writing and

# Corpus Linguistics for Academic Literacies:

---

- 1) Analysis of academic genres
- 2) Analysis of disciplinary differences
- 3) Development of teaching materials
- 4) Evaluation of language proficiency



# What is a corpus?

	HELP	TRANSLATE	ANALYZE
Randi Rippen, Introduction to The Cambridge Handbook Of	Corpus Linguistics	1	( Douglas Biber & Randi Rippen eds. , 2015
at 564. # 24. Paul Baker , Glossary of	Corpus Linguistics	65	) . # 25 . See Tony McEnery &
tool of a linguistic corpus-what we consider the heart of	corpus linguistics	analysis-is	the concordance line,40 Concordance lines
have an interest in reliable measuring instruments are	corpus linguistics	and applied linguistics	( Barnden , Lee , and Markert ;
important for the application of our instrument to	corpus linguistics	and experimental psychology	. A proposal for handling
First , though , we provide a brief introduction to	corpus linguistics	and the databases we analyze . # Corpus linguistics m	
be affected in three major ways by the expansion of	Corpus Linguistics	as a subfield of Applied Linguistics : (1) the growth	
, Using Corpora in the Language Learning Classroom :	Corpus Linguistics	for Teachers	( 2010 ) . # 37 . Id. ;
of language as it is actually used . " 24	Corpus linguistics	gets its name from	the databases (or bodies ) of texts
of large bodies of texts ) has burgeoned . As	corpus linguistics	has grown as	a field , it has greatly impacted both gra
Mascott used a research technique called	corpus linguistics	in addition to	canvassing the historical record , allowir
precedent . # A. The Purpose of Corpus Linguistics #	Corpus linguistics	is an empirical study	of language that is based on the
to represent the speech community they seek to study,25 #	Corpus linguistics	is founded on	the twin ideas that a corpus of texts car
feasible . " 28 Because " a key goal of	corpus linguistics	is to aim for	replicability of results , data creators have
to corpus linguistics and the databases we analyze ; #	Corpus linguistics	may sound enigmatic to	the legal ear , but it has very
on both quantitative and qualitative analysis . " 27 And	corpus linguistics	results in research	findings that have much greater
and grammar research . In her article , " Will	corpus linguistics	revolutionize grammar teaching in	the 21st century ? "
94-112. # Conrad , S. ( 2000 ) : Will	Corpus Linguistics	revolutionize grammar teaching in	the 21st century ? "
mission statements . In the present analysis , however ,	corpus linguistics	software permitted a	comparison of large data sets . 1
since this is what they are referred to in the	corpus linguistics	software used in	this study . ) For example , " Do
, corpus-assisted research design , I used specialized	corpus linguistics	software ( Scott , 2013 )	to compare two sets of textua
institutional environment.Method:Using	corpus linguistics	software I compared	1,009 mission statements from
be present in the main corpus ; I used specialized	corpus linguistics	software Wordsmith Tools 6.0	( Scott , 2013 ) , to
of the study.AnalysisData analysis involved	corpus linguistics	techniques in which	researchers use statistical algorit

# What is a corpus?

---

- A corpus is a large, principled collection of natural texts
- A corpus is a collection of electronic texts usually stored on a computer
- A corpus is available for qualitative and quantitative analysis

(Biber, Conrad and Reppen, 1998; O'Keeffe, McCarthy and Carter, 2003)

# What is a corpus?

---

## (BAWE) British Academic Written English Corpus

### About BAWE

The British Academic Written English Corpus (BAWE) was collected as part of the project, 'An Investigation of Genres of Assessed Writing in British Higher Education'. The project was funded by the [Economic and Social Research Council](#). (2004 - 2007 project number RES-000-23-0800).





# What is a corpus?



List Chart Word E

## (BAWE) British Academic Written English Corpus

### About BAWE

The British Academic Written English Corpus (BAWE) was collected as part of the project, 'An Investigation of Genres of Assessed Writing in British Higher Education'. The project was funded by the [Economic and Social Research Council](#). (2004 - 2007 project number RES-000-23-0800).



# What is a corpus?



## (BAWE) British Academic Written English Corpus

### What is BAWE

The British Academic Written English Corpus (BAWE) was collected as part of a project, 'An Investigation of Genres of Assessed Writing in British Higher Education'. The project was funded by the [Economic and Social Research Council](#) (2004 - 2007 project number RES-000-23-0800).



BRITISH  
NATIONAL  
CORPUS  
2014

# What is a corpus?

**MICASE**

Michigan Corpus of Academic Spoken English



Corpus of

SEARCH

[Home](#)

[Search](#)

[Browse](#)

[Help](#)

30 transcripts

Transcript ID  
(click to view)

File Name

Record

[LEL115SU005](#)

*Medical Anthropology Lecture*

6

[LEL175MU014](#)

*Intro Biology First Day Lecture*

4

[LEL300SU020](#)

*Literature and Social Change Lecture*

8

[LEL500JU034](#)

*Intro Psychology Lecture*

4

[LEL295JU035](#)

*Intro Engineering Lecture*

5

[LEL280JG051](#)

*Graduate Macroeconomics Lecture*

7

[LEL565SU064](#)

*Principles in Sociology Lecture*

8

[LEL185SU066](#)

*Behavior Theory Management Lecture*

11 8

[LEL220JU071](#)

*Intro Communication Lecture*

7

**B N**  
**BRITISH  
NATIONAL  
CORPUS  
2014**

# **What can we do with corpora?**

## **Language Analysis**

---

# What can we do with corpora?

## Language Analysis

Since the compilation of the first corpora, CL has been primarily used to conduct language analysis.

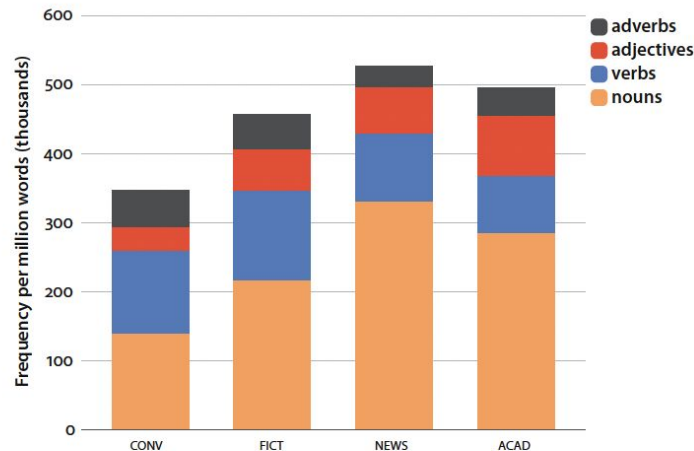
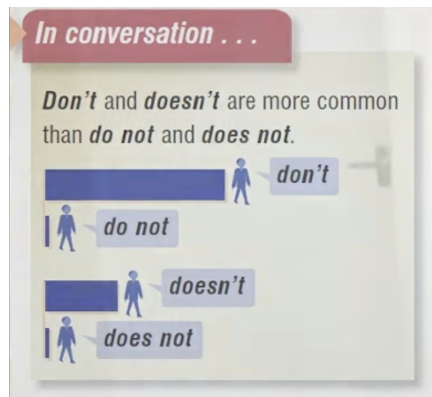


Figure 2.2 Distribution of lexical word classes across registers

# What can we do with corpora?

## Language Analysis

---



# What can we do with corpora? Language Analysis

---



What are the most frequent verbs of academic writing?

# Most frequent verbs in Academic Writing

HELP	i	★		ALL	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
1	i	★	[USE]	211021	211021	27212	32245	32132	35193	43406	40833
2	i	★	[MAKE]	175561	175561	31144	32324	31512	31018	27088	22475
3	i	★	[SEE]	151725	151725	24101	26858	19702	22514	20366	38184
4	i	★	[PROVIDE]	112332	112332	17978	18428	18209	19182	21774	16761
5	i	★	[FIND]	110937	110937	17176	18210	18695	21123	19018	16715
6	i	★	[TAKE]	109983	109983	19301	20212	19388	20135	16751	14196
7	i	★	[SAY]	97418	97418	14611	18739	19214	21877	13592	9385
8	i	★	[BECOME]	90184	90184	18036	17735	16941	15318	11975	10179
9	i	★	[INCLUDE]	85030	85030	12330	13348	14017	15098	16802	13435
10	i	★	[GIVE]	84978	84978	16507	15835	14911	14425	11969	11331
11	i	★	[SHOW]	82536	82536	11579	12684	12988	13507	14168	17610
12	i	★	[KNOW]	71516	71516	12526	12656	12521	14366	11272	8175
13	i	★	[NEED]	66303	66303	9896	10549	11111	12972	13996	7779
14	i	★	[COME]	64484	64484	11884	12395	11687	12480	8877	7161
15	i	★	[WORK]	64445	64445	9610	10266	11406	13536	12746	6881













# Does it vary by discipline?

HELP	i	★		ALL	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
1	i	★	[USE]	211021	<b>211021</b>	27212	32245	32132	35193	43406	40833
2	i	★	[MAKE]	175561	<b>175561</b>	31144	32324	31512	31018	27088	22475
3	i	★	[SEE]	151725	<b>151725</b>	24101	26858	19702	22514	20366	38184
4	i	★	[PROVIDE]	112332	<b>112332</b>	17978	18428	18209	19182	21774	16761
5	i	★	[FIND]	110937	<b>110937</b>	17176	18210	18695	21123	19018	16715
6	i	★	[TAKE]	109983	<b>109983</b>	19301	20212	19388	20135	16751	14196
7	i	★	[SAY]	97418	<b>97418</b>	14611	18739	19214	21877	13592	9385
8	i	★	[BECOME]	90184	<b>90184</b>	18036	17735	16941	15318	11975	10179
9	i	★	[INCLUDE]	85030	<b>85030</b>	12330	13348	14017	15098	16802	13435
10	i	★	[GIVE]	84978	<b>84978</b>	16507	15835	14911	14425	11969	11331
11	i	★	[SHOW]	82536	<b>82536</b>	11579	12684	12988	13507	14168	17610
12	i	★	[KNOW]	71516	<b>71516</b>	12526	12656	12521	14366	11272	8175
13	i	★	[NEED]	66303	<b>66303</b>	9896	10549	11111	12972	13996	7779
14	i	★	[COME]	64484	<b>64484</b>	11884	12395	11687	12480	8877	7161
15	i	★	[WORK]	64445	<b>64445</b>	9610	10266	11406	13536	12746	6881

# Does it vary by discipline?

YES!

SECTION	History	Education	Geog/SocSci	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc	Business
FREQ	17137	60593	58330	18482	27347	11841	62202	40431	8209	3220
WORDS (M)	13.4	15.8	20.0	12.3	16.2	7.8	17.5	10.8	4.8	1.2
PER MIL	1,279.32	3,841.11	2,912.35	1,504.35	1,687.77	1,510.05	3,563.74	3,740.31	1,703.93	2,728.36
CLICK FOR CONTEXT										

# **What can we do with corpora?**

## **Language Teaching**

---

# Corpus Revolution in the Classroom

---

In the 90s, linguists (and lexicographers) started to reflect on how corpus linguistics would revolutionize language teaching.

# Corpus Revolution in the Classroom

---

In the 90s, linguists (and lexicographers) started to reflect on how corpus linguistics would revolutionize language teaching.

---

## The corpus revolution

MICHAEL RUNDELL  
PENNY STOCK



*A consideration of the practical benefits  
to English-language lexicographers of the  
evidence derived from computer corpora  
(second article of three)*

1992

---

# Corpus Revolution in the Classroom

---

In the 90s, linguists (and lexicographers) started to reflect on how corpus linguistics would revolutionize language teaching.

---

## The corpus revolution

MICHAEL RUNDELL  
PENNY STOCK



*A consideration of the practical benefits to English-language lexicographers of the evidence derived from computer corpora (second article of three)*

1992

---

---

## The corpus revolution revisited

MICHAEL RUNDELL

An update on the rise and rise of electronic language corpora and their impact on dictionaries

2008

---

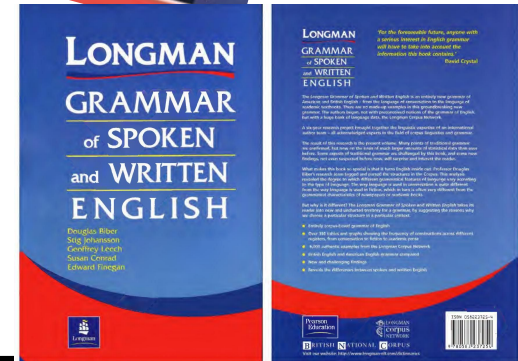
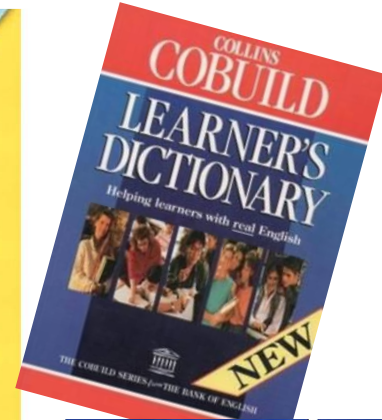
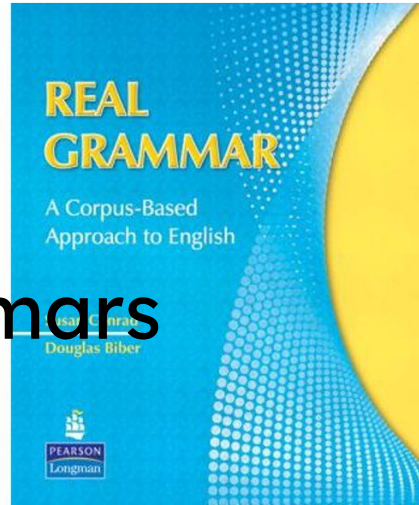
# Corpus Linguistics in Teaching Materials

---

- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment



## The Academic Word List

The Academic Word List is a useful English resource for lecturers and students.

Averil Coxhead from the [School of Linguistics and Applied Language Studies](#) developed

BRIDGE  
GRAMMAR  
AND BEYOND



# Corpus Linguistics in Teaching Materials

---

- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

But, these are indirect applications of corpus to teaching.



To join the **Corpus Revolution!**

# **Why include corpus linguistics in language teaching?**

---

# Why include corpus linguistics in language teaching?

---

## Corpus Use in Language Learning: A Meta-Analysis

Alex Boulton and Tom Cobb

Université de Lorraine and Université du Québec à Montréal

This study applied systematic meta-analytic procedures to summarize findings from experimental and quasi-experimental investigations into the effectiveness of using the tools and techniques of corpus linguistics for second language learning or use, here referred to as data-driven learning (DDL). Analysis of 64 separate studies representing 88 unique samples reporting sufficient data indicated that DDL approaches result in large overall effects for both control/experimental group comparisons ( $d = 0.95$ ) and for pre/posttest designs ( $d = 1.50$ ). Further investigation of moderator variables revealed that small effect sizes were generally tied to small sample sizes. Research has barely begun in some key areas, and durability/transfer of learning through delayed posttesting remains an area in need of further investigation. Although DDL research demonstrably improved over the period investigated, further changes in practice and reporting are recommended.

64 studies on the applications  
of Data-Driven Learning  
(Johns, 1990)

# Corpus Revolution in the Classroom

---

## Towards the corpus revolution? Bridging the research–practice gap

Angela Chambers

University of Limerick, Ireland

Email: [Angela.Chambers@ul.ie](mailto:Angela.Chambers@ul.ie)

Although there is a large and increasing body of research on the use of corpus data by language teachers and learners, **the language teachers in question are also the researchers reporting on their use of corpora with their own learners.**

# Why include corpus linguistics in language teaching?

---

- Develops students' autonomy as language learners
- Identifies language patterns (e.g., frequency, co-occurrence patterns, etc), rather than rules (Boulton & Cobb, 2017)
- Uses authentic language data
- Helps learners understand conventions of specific discourse communities

# Why include corpus linguistics in language teaching?

---

- Develops students' autonomy as language learners
- Identifies language patterns (e.g., frequency, co-occurrence patterns, etc), rather than rules (Boulton & Cobb, 2017)
- Uses authentic language data
- Helps learners understand conventions of specific discourse communities

**Are multi-word verbs (e.g., *point out*, *look at*) more frequent in spoken or written texts?**

---





**Are multi-word verbs (e.g., *point out*, *look at*) more frequent in spoken or written texts?**

---

***point out***

# Are multi-word verbs (e.g., *point out*, *look at*) more frequent in spoken or written texts?

## *point out*

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
FREQ	19127	4471	3311	482	4440	818	1882	1793	1930
WORDS (M)	993	20.7	20.1	20.7	20.4	19.1	20.4	19.6	19.3
PER MIL	19.26	214.79	164.65	23.25	217.49	42.72	92.22	91.00	99.55
SEE ALL SUB-SECTIONS AT ONCE									

**Are multi-word verbs (e.g., *point out*, *look at*) more frequent in spoken or written texts?**

---

***look at***

# Are multi-word verbs (e.g., *point out*, *look at*) more frequent in spoken or written texts?

***look at***

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
FREQ	235785	27128	22643	58463	62809	26223	15745	16088	6686
WORDS (M)	993	20.7	20.1	20.7	20.4	19.1	20.4	19.6	19.3
PER MIL	237.44	1,303.27	1,125.97	2,820.46	3,076.70	1,369.36	771.54	816.51	344.86
SEE ALL SUB-SECTIONS AT ONCE									

**How can teachers bring corpus tools to the classroom?**

---

# Online Corpus Tools

# Linggle

discuss ?about the n.



Sketch Engine for language learning

word or phrase



English



flax

interactive language learning

## Learning Collocations

claim in academic English in Social Sciences go

Your IP

Click to show useful academic words

Family Words | Synonyms | Antonyms

claimed claiming claims unclaimed

claim used as a noun | claim used as a verb

adjective + claim	second <b>claim</b>	12	first <b>claim</b>
	latter <b>claim</b>	5	factual <b>claim</b>
	general <b>claim</b>	3	earlier <b>claim</b>
	average <b>claim</b>	3	main <b>claim</b>

# Online Corpus Tools

---

**CORPUS**  
**MATE**

Enter your search here

in topic



in mode



Search

Show results as  Pattern view  Sentences

*jitw*

Enter a word or short phrase:

combinations

alternatives from thesaurus

alternatives from learner errors



# **10 ideas of activities that bring corpus tools and techniques to the English classroom!**

---



# Activity #1: Collocates of Academic Words

---



In this essay, I provide the results of a study on the use of computers in the classroom.

# Activity #1: Collocates of Academic Words

---

- a) Select the vocabulary you would like to practice with your students

*e.g., claim, argument, existence, evidence*

## The Academic Word List

Headword	Sublist	Related word forms
analyse	1	analysed, analyser, analysers, analytical, analytically, analyze,
approach	1	approachable, approached, ap
area	1	areas
assess	1	assessable, assessed, assesses, reassessed, reassessing, reass

# Activity #1: Collocates of Academic Words

---

## b) Search their collocates in Flax

### *Learning Collocations*



in

Your ID

#### **λ About this collection**

There are many definitions of collocation. We think of collocations in the same way as expressed by Benson et al.: "In any language, certain words occur with certain other words or grammatical constructions. These recurrent, semi-fixed combinations, or collocations, can be divided into two groups: grammatical and lexical collocations."

Enter this collection to search for the company words keep. How words form into collocational patterns will be revealed by looking across the 100 National Corpus (BNC) reference corpus, the British Academic Written English (BAWE) corpus of 2500 university student writing texts and the Wiki corpus of three million articles.



# Activity #1: Collocates of Academic Words

## b) Search their collocates in Flax

### Learning Collocations

claim in contemporary English (Wikipedia) go Y

Click to show useful academic words

Family Words | Synonyms | Antonyms

*claimed claiming claims unclaimed*

claim used as a verb | claim used as a noun

verb + claim	 lay <b>claim</b> to	488	 support this <b>claim</b>	145
	 denied the <b>claim</b>	63	 renounced his <b>claim</b> to	46
	 disputed this <b>claim</b>	36	 make this <b>claim</b>	31
	 dismissed the <b>claim</b>	30	 rejected this <b>claim</b>	26
	 prove his <b>claim</b>	23	 did <b>claim</b>	21

>>> more

# Activity #1: Collocates of Academic Words

---

c) Create a matching exercise! (or a memory game)

a claim

based

the existence

provide

evidence

this argument

confirmed

support

## Activity #2: Key Word In Context (KWIC)

---

ESL students often struggle with the difference between “through” and “toward”

I walked through the door?

I walked toward the door?

# Activity #2: Key Word In Context (KWIC)

---

a) Search the words in SKELL



*Sketch Engine for language learning*

Search bar with a magnifying glass icon, the text "word or phrase", a keyboard icon, and a language dropdown menu set to "English".

ABOUT SKELL



<https://skell.sketchengine.eu/>

# Activity #2: Key Word In Context (KWIC)

## b) Ask students to explain the difference between the two in SKELL

through 751.08 hits per million



1. When bank owned homes are sold **through** bank loss mitigation.
2. This section comprises chapters 12 **through** 17.
3. He finished high school **through** independent study.
4. I achieved block captain status **through** default.
5. The program is administered **through** existing agencies.
6. This balancing act is achieved **through** measurement.
7. The something is associated discount cialis australia **through** uncontrollable attaining.
8. The benefits offer smart climate control **through** dynamic heat insulation.
9. **Through** international agreements data are exchanged between countries worldwide.
10. Every field is going **through** rapid change.
11. A search **through** many alternative solutions is often required.
12. The current analysis includes data **through** 2005.
13. The war progressed **through** three distinct stages.
14. Successful interventions are subsequently promoted **through** training.

toward 73.03 hits per million

1. There are several important steps **toward** such liberation.
2. The progressive attitude **toward** history is sufficiently revealing.
3. You are going forward **toward** something great.
4. Competition eventually shifts temporarily high corporate profitability **toward** lower
5. The journey **toward** making "normal" obsolete.
6. This condition may proceed **toward** younger leaves.
7. The official view **toward** poker resistance here.
8. We were definitely more aggressive **toward** batters.
9. These reports are oriented **toward** total program evaluation.
10. Market structures oriented **toward** accommodating supply side resources 13.
11. The push **toward** mandatory policies gathered momentum more recently.
12. The activities are directed **toward** family participation.





## Activity #2: Key Word In Context (KWIC)

---

While the words “through” and “toward” have different meanings, they appear in similar contexts. For each of the following sentences, fill in the blank with the correct word: “through” or “toward.”

- 1) Knife blades should be turned \_\_\_\_\_ the plate. (WEB)
- 2) You walk \_\_\_\_\_ that door, you’re not coming back. (BLOG)
- 3) Gardner has been working \_\_\_\_\_ this goal for most of her life. (BLOG)
- 4) Government loans and grants helped me \_\_\_\_\_ school. (BLOG)
- 5) Mbasi turned and looked \_\_\_\_\_ the captain. (FIC)
- 6) I’m so glad you all made it \_\_\_\_\_ that awful storm. (BLOG)
- 7) The sun came \_\_\_\_\_ the leaves and there was a little breeze stirring. (FIC)
- 8) I’ve had a lot of interns pass \_\_\_\_\_ my office. (BLOG)

**We can also use corpus research to inform material development**

---

## Activity #3: Suffixes (nominalization)

---

Which derivational suffixes are the most frequent in academic writing?

- ism (capitalism)
- tion (operation)
- ity (flexibility)
- ness (happiness)
- ure (exposure)
- ance (governance)



## Activity #3: Suffixes (nominalization)

---

Which derivational suffixes are the most frequent in academic writing?

- a) Give students a text and ask them to:
  - i) highlight all the nouns
  - ii) Underline the suffixes that are used to form each noun
  - iii) Count the frequency of suffixes

## Activity #3: Suffixes (nominalization)

---

Corporate social responsibility has become a widely held buzzword in the business community for good corporate practices. However, it has also been the focus of debate in its very nature and theoretical justification. One of these criticisms is that corporate social responsibility sounds fine in theory but is unworkable in practice. This article investigates the validity of this claim. First by briefly looking at the nature of corporate social responsibility.

**(BAWE) British Academic Written English Corpus**

# Activity #4: Understanding the differences between spoken and written discourse

---

Research shows that:

- Spoken = clausal
- Written = phrasal

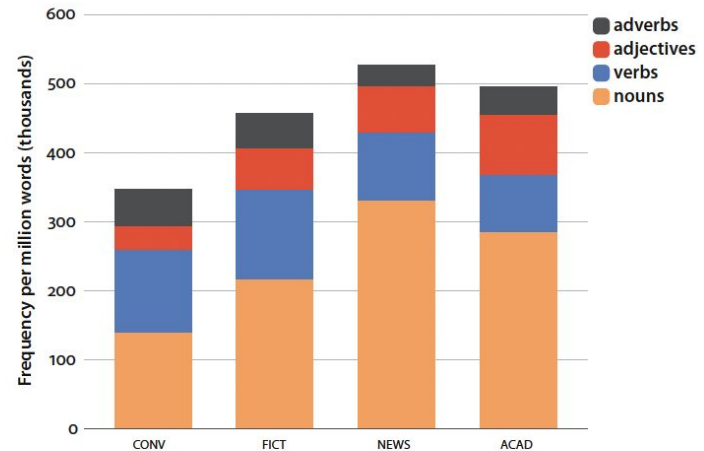


Figure 2.2 Distribution of lexical word classes across registers

## **Activity #4: Understanding the differences between spoken and written discourse**

---

- a) Give students an excerpt from an academic text and a spoken text and ask them to:
  - i) highlight in color 1 all the nouns
  - ii) highlight in color 2 all the verbs
  - iii) Compare the frequency of both



# Activity #4: Understanding the differences between spoken and written discourse

---

Corporate social responsibility has become a widely held buzzword in the business community for good corporate practices. However, it has also been the focus of debate in its very nature and theoretical justification. One of these criticisms is that corporate social responsibility sounds fine in theory but is unworkable in practice. This article investigates the validity of this claim.

(BAWE) British Academic Written English Corpus

I think the intention is just to see if, people use big words or little in the classroom and what kind of words they use, and to once they know that that should help them, uh train T-As, train G-S-Is help people speak English as a second language. so we'll see what comes out of it maybe i'll speak good today. <SS: LAUGH> um, okay... so let me get to my notes, so i've already introduced myself i've been at the university three years i came here from the University of Washington.



Michigan Corpus of Academic Spoken English

# Activity #4: Understanding the differences between spoken and written discourse

---

Corporate social responsibility has become a widely held buzzword in the business community for good corporate practices. However, it has also been the focus of debate in its very nature and theoretical justification. One of these criticisms is that corporate social responsibility sounds fine in theory but is unworkable in practice. This article investigates the validity of this claim.

(BAWE) British Academic Written English Corpus

I think the intention is just to see if, people use big words or little in the classroom and what kind of words they use, and to once they know that that should help them, uh train T-As, train G-S-Is help people speak English as a second language. so we'll see what comes out of it maybe i'll speak good today. <SS: LAUGH> um, okay... so let me get to my notes, so i've already introduced myself i've been at the university three years i came here from the University of Washington.



Michigan Corpus of Academic Spoken English

# Hands-on vs Hands-off Activities

---

**Hands-off:** Pre-prepared activities by the teacher (e.g., handouts)

**Hands-on:** Students use corpus tools in the classroom

## Activity #5: Comparing synonyms

---

Attributive adjectives are characteristics of academic writing.

e.g.,     a **critical** argument...  
          an **important** theory...  
          a **respectable** claim...

## Activity #5: Comparing synonyms

---

- a) Ask students about two adjectives that they think are synonymous.
- b) Ask students to make a list of nouns that appear right after the selected adjectives (*e.g., important, critical, fundamental, etc*)

# Activity #5: Comparing synonyms

---

For example:

fundamental  
pivotal

+ \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_  
+ \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_

## Activity #5: Comparing synonyms

---

For example:

fundamental  
pivotal

+ *theory, resource, study*  
+ *theory, study, finding*

# Activity #5: Comparing synonyms

---

d) compare the collocates of both adjectives using the compare function on COCA

List Chart Word Browse Collocates **Compare** KWIC -

fundamental Word1 adj.ALL  ▶\_pos

pivotal Word2 adj.ALL

\* Collocates [POS]

+	4	3	2	1	0	0	1	2	3	4	+
---	---	---	---	---	---	---	---	---	---	---	---

Sections Texts/Virtual Sort/Limit Options



# Activity #5: Comparing synonyms

d) compare the collocates of both adjectives using the compare function on COCA

WORD 1 (W1): FUNDAMENTAL (9.52)

	WORD	W1	W2	W1/W2	SCORE
1	RIGHTS	251	0	502.0	52.7
2	CHANGES	163	0	326.0	34.2
3	PRINCIPLES	162	0	324.0	34.0
4	PROBLEM	145	0	290.0	30.5
5	RIGHT	141	0	282.0	29.6
6	PRINCIPLE	125	0	250.0	26.3
7	HUMAN	121	0	242.0	25.4
8	DIFFERENCES	104	0	208.0	21.8
9	SHIFT	86	0	172.0	18.1
10	FREQUENCY	79	0	158.0	16.6
11	VALUES	78	0	156.0	16.4
12	RESEARCH	73	0	146.0	15.3
13	FREEDOMS	71	0	142.0	14.9
14	PROBLEMS	69	0	138.0	14.5
15	WAY	64	0	128.0	13.4
16	MOVEMENT	59	0	118.0	12.4
17	DIFFERENCE	108	1	108.0	11.3
18	UNDERSTANDING	53	0	106.0	11.1
19	OPTION	51	0	102.0	10.7

WORD 2 (W2): PIVOTAL (0.11)

	WORD	W2	W1	W2/W1	SCORE
1	MOMENT	41	1	41.0	390.4
2	MOMENTS	10	0	20.0	190.4
3	BEHAVIOR	10	0	20.0	190.4
4	EVENT	17	1	17.0	161.9
5	FIGURE	16	1	16.0	152.4
6	STATES	6	0	12.0	114.3
7	EVENTS	10	1	10.0	95.2
8	MAN	4	0	8.0	76.2
9	SCENE	4	0	8.0	76.2
10	TIME	7	1	7.0	66.7
11	EPISODE	3	0	6.0	57.1
12	CLINICAL	3	0	6.0	57.1
13	YEARS	3	0	6.0	57.1
14	TARGET	3	0	6.0	57.1
15	SCENES	3	0	6.0	57.1
16	MEMBER	3	0	6.0	57.1
17	PHASE	3	0	6.0	57.1
18	RESPONSE	15	3	5.0	47.6
19	CASE	4	1	4.0	38.1

## Activity #5: Comparing synonyms

---

e) compare the COCA results to the initial answers:

- Are the collocates the same as you wrote in exercise c)?
- Are the collocates of both words similar?
- How would you explain the difference between both words?

# Activity #6: Identifying Collocates

---

Which words collocate with *study*?

jitw

Enter a word or short phrase:

combinations

alternatives from thesaurus

alternatives from learner errors

Recent searches

# Activity #6: Identifying Collocates

## a) Search for *study* in Just-the-word

### study (N)

[V obj \\*study\\*](#), e.g. undertake study  
[\\*study\\* subj V](#), e.g. study show  
[\\*study\\* subj N](#), e.g. study study  
[ADJ \\*study\\*](#), e.g. recent study  
[N \\*study\\*](#), e.g. case study  
[\\*study\\* PREP](#), e.g. study of  
[\\*study\\* V](#), e.g. study using  
[\\*study\\* V](#), e.g. study conducted  
[\\*study\\* N](#), e.g. study group  
[N N \\*study\\*](#), e.g. child development study  
[PREP \\*study\\*](#), e.g. in study  
[N PREP \\*study\\*](#), e.g. result of study  
[ADJ PREP \\*study\\*](#), e.g. available for study  
[N and \\*study\\*](#), e.g. work and study  
[\\*study\\* and N](#), e.g. study and study  
[article \\*study\\*](#), e.g. studies

### study (V)

[\\*study\\* obj N](#), e.g. study subject  
[N subj \\*study\\*](#), e.g. student study  
[ADV \\*study\\*](#), e.g. also study  
[\\*study\\* ADV](#), e.g. study carefully  
[\\*study\\* PREP](#), e.g. study at  
[V and \\*study\\*](#), e.g. look and study  
[\\*study\\* and V](#), e.g. study and find

*phrase (nn)* nn is the frequency of usage

— Good Word Combinations

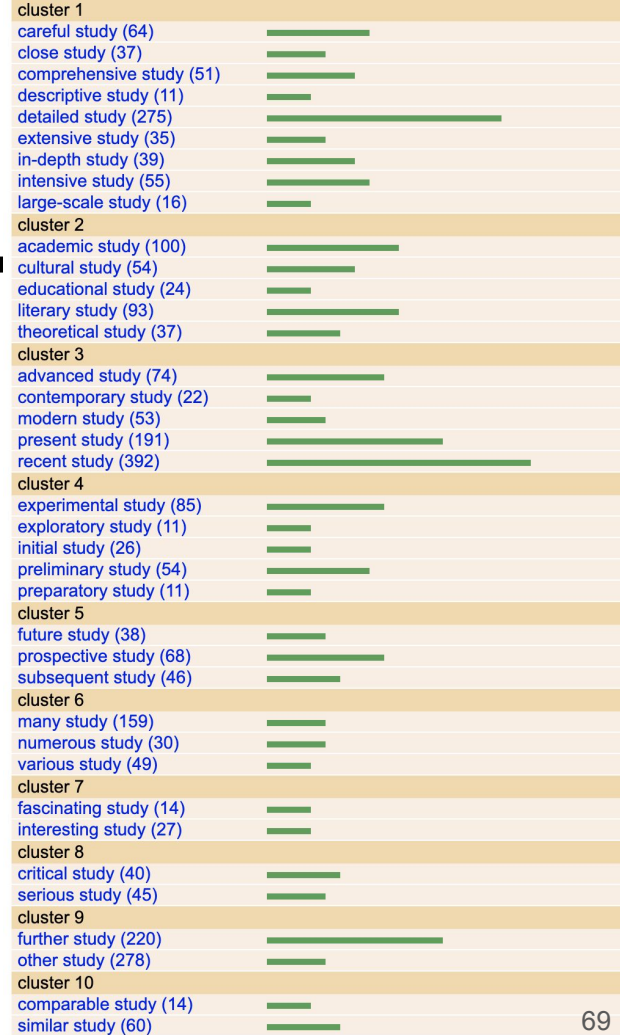
— Bad Word Combinations

— Similarity of meaning

.word - Means singular noun only

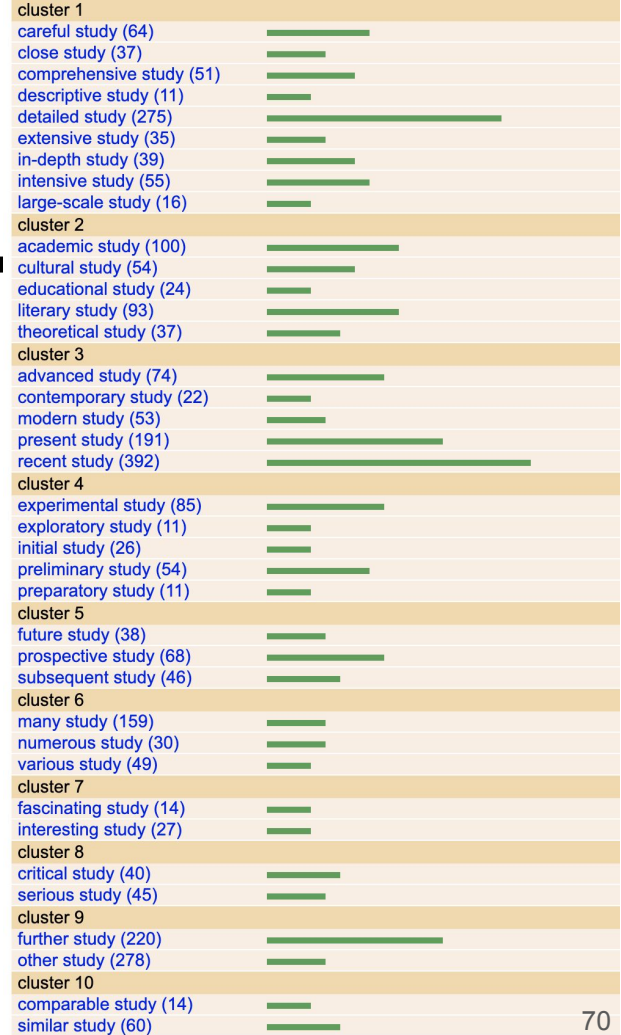
# Activity #6: Identifying Collocates

- Search for *study* in Just-the-word
- Which adjectives are used to describe *study* in JTW?



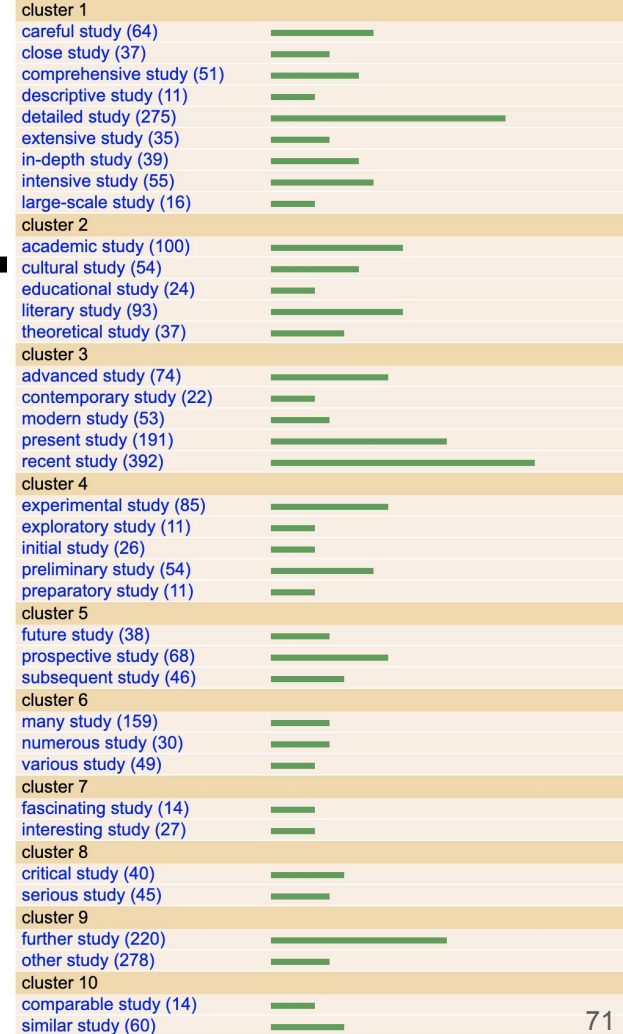
# Activity #6: Identifying Collocates

- Search for *study* in Just-the-word
- Which adjectives are used to describe *study* in JTW?
- Try the same search with different words (e.g., *research*, *investigation*).



# Activity #6: Identifying Collocates

- Search for *study* in Just-the-word
- Which adjectives are used to describe *study* in JTW?
- Try the same search with different words (e.g., *research*, *investigation*).
- Explain how these words are used in academic English.



## **Activity #7: Using corpus tools to analyze a text**

---

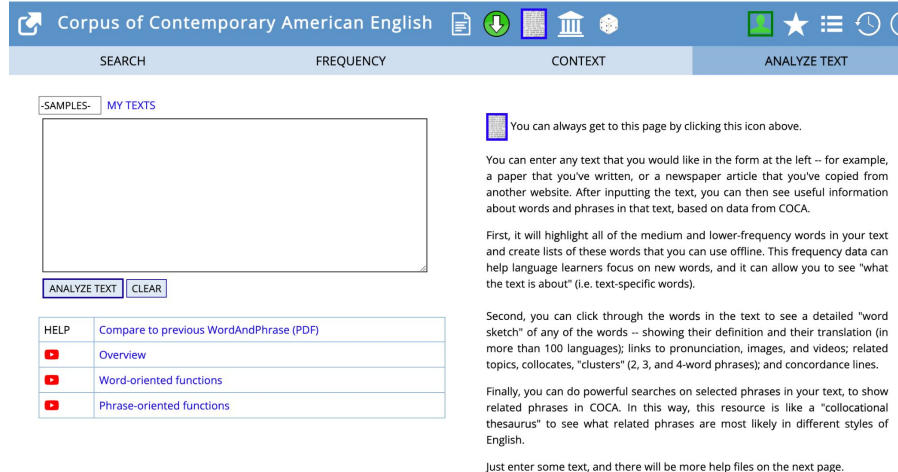
There are two tools teachers can use to help learners analyze their own texts.



# Activity #7: Using corpus tools to analyze a text

There are two tools teachers can use to help learners analyze their own texts.

## Word and Phrase






The screenshot shows the COCA Word and Phrase tool interface. At the top, there is a navigation bar with tabs for SEARCH, FREQUENCY, CONTEXT, and ANALYZE TEXT. The ANALYZE TEXT tab is currently selected. Below the navigation bar, there is a large text input area with a placeholder box. To the right of the input area, there is a help icon and a text box that reads: "You can always get to this page by clicking this icon above." Below the input area, there are buttons for ANALYZE TEXT and CLEAR. At the bottom, there is a table with a HELP column and a list of links: "Compare to previous WordAndPhrase (PDF)", "Overview", "Word-oriented functions", and "Phrase-oriented functions".

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT ANALYZE TEXT

-SAMPLES- MY TEXTS

ANALYZE TEXT CLEAR

HELP	Compare to previous WordAndPhrase (PDF)
	<a href="#">Overview</a>
	<a href="#">Word-oriented functions</a>
	<a href="#">Phrase-oriented functions</a>

You can always get to this page by clicking this icon above.

You can enter any text that you would like in the form at the left – for example, a paper that you've written, or a newspaper article that you've copied from another website. After inputting the text, you can then see useful information about words and phrases in that text, based on data from COCA.

First, it will highlight all of the medium and lower-frequency words in your text and create lists of these words that you can use offline. This frequency data can help language learners focus on new words, and it can allow you to see "what the text is about" (i.e. text-specific words).

Second, you can click through the words in the text to see a detailed "word sketch" of any of the words – showing their definition and their translation (in more than 100 languages); links to pronunciation, images, and videos; related topics, collocates, "clusters" (2, 3, and 4-word phrases); and concordance lines.

Finally, you can do powerful searches on selected phrases in your text, to show related phrases in COCA. In this way, this resource is like a "collocational thesaurus" to see what related phrases are most likely in different styles of English.

Just enter some text, and there will be more help files on the next page.

# Activity #7: Using corpus tools to analyze a text

EDIT TEXT	SAVE TEXT	<input checked="" type="radio"/> WORD	<input type="radio"/> PHRASE
FREQ RANGE	1-500	501-3000	> 3000
290 WORDS	47 %	13 %	19 %

CLICK ON ANY WORD BELOW FOR A FULL WORD SKETCH

The **fish-tetrapod transition** has been called the **greatest** step in **vertebrate** history (Long and Gordon, 2004) and even one of the most **significant events** in the history of life (Carroll, 2001). **Indeed**, the **morphological**, **physiological**, and **behavioral** changes **necessary** for such a **transformation in lifestyle** to **occur** are **astounding**. The **sum** of these **modifications occurring** during the Devonian and Carboniferous led to the **eventual filling** of the **terrestrial realm** with **vertebrate** life, **forever altering** the **structure** and **ecology** of **terrestrial communities**. Long and Gordon (2004) **cited** **six** **critical** questions **relating** to the **evolution of tetrapods**. These questions aimed to **ascertain** which **sarcopterygian** fish were **basal to tetrapods**, **how morphological** changes **occurred sequentially**, and **when, where, how, and why** these changes took place. Many **researchers** have **described** the **morphological** changes that **occurred** (Clack, 2002b; Eaton, 1951; Jarvik, 1955; Long and Gordon, 2004; Thomson, 1993), and others have focused **specifically** on the development of **limbs** and **digits** (Clack, 2002b; Coates and Clack, 1990; Coates et al, 2002; Daeschler and Shubin, 1995; Shubin et al, 1997; Shubin et al, 2004). As Long and Gordon (2004) pointed out, the question that is the least well answered is the question of why these **modifications** occurred. **Exactly** what **factors** drove these changes to take place? Many **researchers** have **posited** theories over the years **attempting** to answer this question, and the **aim** of this **paper** is to **assess** these **arguments** and suggest some **possible common** causes that could **tie** many of the **proposed causal** factors together. However, a **brief description** of known **data** **pertaining** to the time and place of **tetrapod origins** is **first necessary** in order to make **valid** statements regarding possible **influential** factors.

(CLICK ANY WORD FOR FULL WORD SKETCH)

LOW FREQ	MID FREQ	HIGH FREQ
3: morphological 2: modifications, terrestrial, tetrapods, vertebrate 1: altering, ascertain, astounding, basal, behavioral, causal, cited, description, digits, ecology, eventual, evolution, fish-tetrapod, forever, influential, lifestyle, limbs, origins, pertaining, physiological, posited, realm, sarcopterygian, sequentially, specifically, sum, tetrapod, transformation, transition, valid	3: al, et, factors, occurred 2: necessary, possible, researchers 1: aim, aimed, arguments, assess, attempting, brief, common, critical, data, described, events, exactly, filling, fish, greatest, indeed, occur, occurring, paper, proposed, regarding, relating, significant, statements, structure, theories, tie	20: the 17: and 12: of 10: to 6: these 5: changes 4: in, is 3: have, many, place, question, that 2: a, history, how, life, questions, this, why 1: answer, answered, are, as, been, called, causes, communities, could, development, drove, during, even, first, focused, for, has, however, known, least, led, make, most, on, one, order, others, out, over, pointed, six, some, step, such, suggest, take, time, together, took, well, were, what, when, where, which, with, years

# Activity #7: Using corpus tools to analyze a text

There are two tools teachers can use to help learners analyze their own texts.

## LexTutor

**Compleat Lexical Tutor v.8.5**  
For data-driven language learning on the Web

QUICK LOOK-UP or LISTEN  
Enter words to search or speak...

BEST SCREEN : Browser Edge/Chrome, Zoom 99/90/100% at 13/14-15/17" [?] No cookies, registration, ads Best of Canada since Trivial Pursuit - Thornbury  
YOUR SCREEN : Browser Chrome Zoom 100% Size 27" IP 73.197.225.15

VP 1-WD Enter a wo >>

QUICKLINKS > FRENCH TOP... TUTORIAL GAMES RESEARCH

- List\_Learn
  - Group Lex<sup>M</sup>
  - Hypertext
  - RA Reading  
Resources-Assisted
  - Corpus Grammar
  - Concord Writer
  - Research<sup>M</sup>
- Multiwords Range
  - PSYCHO-LEX
  - ReacTime<sup>M</sup>
  - STM<sup>M</sup>
  - Vocab Stats
  - Concordance
  - Vocabprofile
- Text Tools
  - Coverage<sup>M</sup>
  - Familizer
  - Frequency  
Nuclear
  - Text Compare
  - RaNdOmICity
  - KeyWords
- Clickers<sup>M</sup>
  - Flashcards<sup>M</sup>
  - Tests<sup>M</sup>
  - Cloze  
Tune-Up Mar 2023
  - Fon(W)Fs  
Hard Spell  
Dictator, ID-Word
  - MorphoLex<sup>M</sup>

# Activity #7: Using corpus tools to analyze a text

	Families	Types	Tokens	Percent
<b>K1 Words (1-1000):</b>	73	83	188	67.38%
Function:	...	...	(106)	(37.99%)
Content:	...	...	(82)	(29.39%)
> Anglo-Sax	...	...	(27)	(9.68%)
<b>K2 Words (1001-2000):</b>	9	10	10	3.58%
> Anglo-Sax	...	...	(4)	(1.43%)
1k+2k			...	(70.96%)
<b>AWL Words:</b>	22	24	30	10.75%
> Anglo-Sax	...	...	()	(0.00%)
<b>Off-List Words:</b>	?	32	51	18.28%
	104+?	149	279	100%

Current profile		
%	Cumul.	
67.38	67.38	Words in text (tokens): 279
		Different words (types): 149
		Type-token ratio: 0.53
		Tokens per type: 1.87
		Lex density (content words/total): 0.62
		<i>Pertaining to onlist only</i>
		Tokens: 228
		Types: 117
		Families: 104
		Tokens per family: 2.19
		Types per family: 1.12
		Anglo-Sax Index: %
		(A-Sax tokens + functors / onlist tokens)
		Greco-Lat/Fr-Cognate Index: (inverse of above) %

ransition has been called the greatest step in vertebrate jordon number and even one of the most significant ory of life carroll number indeed the morphological behavioral changes necessary for such a transformation ar are astounding the sum of these modifications the devonian and carboniferous led to the eventual filling ealm with vertebrate life forever altering the structure and rial communities

number cited six critical questions relating to the evolution of tetrapods these questions aimed to ascertain which sarcopterygian fish were basal to tetrapods how morphological changes occurred sequentially and when where how and why these changes took place many researchers have described the morphological changes that occurred clack numberb eaton number jarvik number long and gordon number thomson number and others have focused specifically on the development of limbs and digits clack numberb coates and clack number coates et al number daeschler and shubin number shubin et al number shubin et al number as long and gordon number pointed out the question that is the least well answered is the question of why these modifications occurred exactly what factors drove these changes to take place many researchers have posited theories over the years attempting to answer this question and the aim of this paper is to assess these arguments and suggest some possible common causes that could tie many of the proposed causal factors together however a brief description of known data pertaining to the time and place of tetrapod origins is first necessary in order to make valid statements regarding possible influential factors

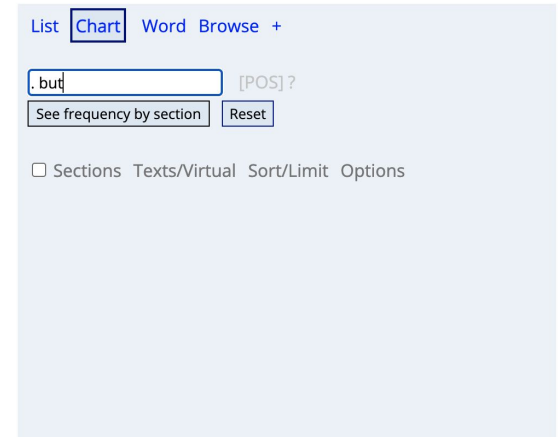
# **Activity #8: And, But, Yet and more**

---

# Activity #8: And, But, Yet and more

---

- a) Compare the frequency of conjunctions in the beginning of a sentence between academic writing and other registers using COCA charts.



The screenshot shows the COCA Word Browse interface. At the top, there are three tabs: "List", "Chart", and "Word Browse +". The "Chart" tab is currently selected. Below the tabs, there is a search input field containing the text ". but" and a "[POS]?" label to its right. Underneath the search field, there are two buttons: "See frequency by section" and "Reset". At the bottom of the interface, there is a checkbox labeled "Sections" followed by the text "Texts/Virtual Sort/Limit Options".

# Activity #8: And, But, Yet and more

- a) Compare the frequency of conjunctions in the beginning of a sentence between academic writing and other registers using COCA charts.

CLICK TO SEE CONTEXT

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD
FREQ	822900	111810	118760	1801	188434	121899	131025	102940	46231
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8
PER MIL	828.68	869.35	955.79	14.06	1,493.90	1,030.23	1,039.13	845.56	385.93
SEE ALL SUB-SECTIONS AT ONCE									

List **Chart** Word Browse +

.but [POS]?

See frequency by section Reset

Sections Texts/Virtual Sort/Limit Options

## **Activity #8: And, But, Yet and more**

---

- b) Ask students to try the same search with different conjunctions (and, yet)
  
- c) Invite students to compare the results for each conjunction and write an explanation for how they are used in academic writing



## **Activity #9: Does discipline matter?**

---

Compare the use of abstract nouns vs concrete nouns.

# Activity #9: Does discipline matter?

---

Compare the use of abstract nouns vs concrete nouns.

**CORPUS**  
**MATE**

Enter your search here in topic in mode Search

Show results as  Pattern view  Sentences

# Activity #9: Does discipline matter?

---

Some ideas of words that could be used in the search:

- Abstract: *idea, theory, hypothesis, explanation, etc*
- Concrete: *instrument, system, equipment, etc*

**CORPUS**  
**MATE**

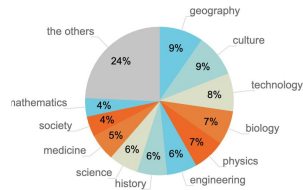
Enter your search here    in topic    in mode    **Search**

Show results as  Pattern view  Sentences

# Activity #9: Does discipline matter?

## Equipment

Topic distribution of documents

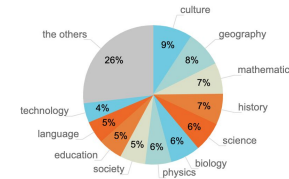


Click on any topic from the list below to see pattern/sentence results only for that topic.

Topic	Documents	%
<a href="#">geography</a>	334	9.68
<a href="#">culture</a>	322	9.34
<a href="#">technology</a>	286	8.29
<a href="#">biology</a>	257	7.45
<a href="#">physics</a>	243	7.05
<a href="#">engineering</a>	228	6.61
<a href="#">history</a>	225	6.52
<a href="#">science</a>	221	6.41
<a href="#">medicine</a>	205	5.94
<a href="#">society</a>	151	4.38
<a href="#">mathematics</a>	139	4.03

## Theory

Topic distribution of documents



Click on any topic from the list below to see pattern/sentence results only for that topic.

Topic	Documents	%
<a href="#">culture</a>	1230	9.18
<a href="#">geography</a>	1128	8.42
<a href="#">mathematics</a>	966	7.21
<a href="#">history</a>	960	7.17
<a href="#">science</a>	929	6.94
<a href="#">biology</a>	923	6.89
<a href="#">physics</a>	813	6.07
<a href="#">society</a>	795	5.94
<a href="#">education</a>	750	5.6
<a href="#">language</a>	696	5.2
<a href="#">technology</a>	605	4.52
<a href="#">medicine</a>	534	3.99
<a href="#">economics</a>	500	3.73
<a href="#">politics</a>	475	3.55
<a href="#">law</a>	440	3.29
<a href="#">psychology</a>	408	3.05
<a href="#">engineering</a>	400	2.99

## **Activity #9: Does discipline matter?**

---

Give students texts from different disciplines and ask them to highlight concrete nouns vs abstract nouns. Then, ask students to write a reflection about the patterns they found in the excerpts.

# Activity #10: However, Nevertheless, Therefore and other linking adverbials

**MICUSP Simple BETA**  
Michigan Corpus of Upper-Level Student Papers

SEARCH CLEAR SEARCH  Include notes & references

You are browsing papers in 16 disciplines at 4 levels of 7 paper types with 8 textual features.

**STUDENT LEVELS**

**NATIVENESS**

**TEXTUAL FEATURES**

**PAPER TYPES**

No Restriction

Argumentative Essay

Creative Writing

Critique/Evaluation

Proposal

Report

Research Paper

Response Paper

**DISTRIBUTION ACROSS DISCIPLINES**

CLICK TO SELECT CLEAR SELECTION

Showing 1 to 20 of 829 papers

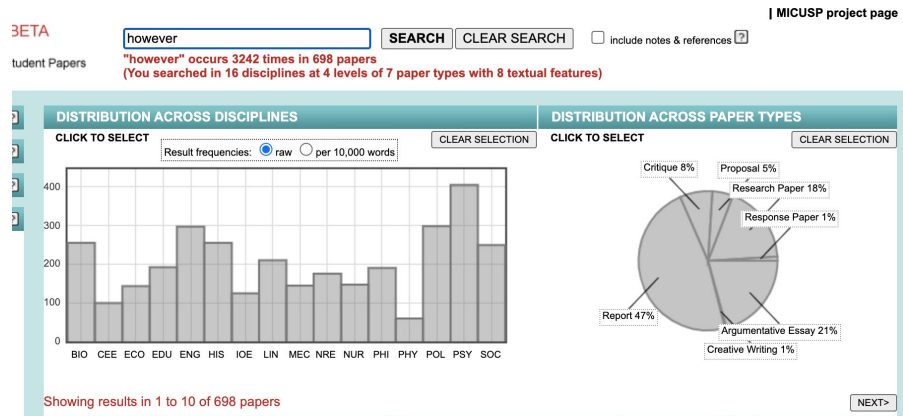
**DISTRIBUTION ACROSS PAPER TYPES**

CLICK TO SELECT CLEAR SELECTION

Paper ID	Title	Discipline	Paper Type
BIO.G0.15.1	Invading the Territory of Invasives: The Dangers of Biotic Disturbance	Biology	Argumentative Essay
BIO.G1.04.1	The Evolution of Terrestriality: A Look at the Factors that Drove Tetrapods to Move Onto Land	Biology	Argumentative Essay
BIO.G3.03.1	Intracellular Electric Field Sensing using Nano-sized Voltmeters	Biology	Argumentative Essay
BIO.G0.11.1	Exploring the Molecular Responses of Arabidopsis in Hypobaric Environments: Identifying Possible Targets for Genetic Engineering	Biology	Proposal
BIO.G1.01.1	V. Cholerae: First Steps towards a Spatially Explicit Model	Biology	Proposal
BIO.G1.07.1	Zebrafish and PGC mis-migration	Biology	Proposal

# Activity #10: *However, Nevertheless, Therefore* and other linking adverbials

Search the MICUSP corpus for one of these linking adverbials and (a) explain where it occurs in the sentence, (b) whether it could occur in other places in the sentence.



<a href="#">BIO.G0.02.4</a>	Chromosomal Rearrangements, Recombination Suppression, and Speciation: A Review of Rieseberg 2001	Biology	Report
	<p>1. From these considerations, Rieseberg offers a more simplistic model to reduce gene flow after a chromosomal rearrangement, than heterospecific fitness. She observes that these rearrangements can strongly effect recombination without rendering the offspring infertile. If the rearrangement is within or linked to an isolating locus of the genome, then reduced gene flow will occur as a result of a lack of recombination. Chromosomal rearrangements can function in this way even when they do not affect hybrid fitness. This theory, <b>however</b>, is still limited because reduced gene flow itself may not cause speciation; furthermore, it is only with an accumulation of many neutral or weakly underdominant rearrangements, linked to isolation loci, that a larger portion of the genome can be affected. One limitation of this model, <b>however</b>, is that it does not help determine whether the arrangement occurred before or after speciation. <b>However</b>, unlike many of the present models, it does allow for speciation in sympatry or parapatry, as well as, reinforcement of allopatrically derived species - which have come together in secondary contact. Rieseberg's new theoretical mechanism proves to be both interesting and plausible. Given further empirical data, it may serve as the new standard model of chromosomal speciation.</p>		
<a href="#">BIO.G0.02.5</a>	On the Origins of Man: Understanding the Last Two Million Years	Biology	Report
1 of 13 hits <a href="#">Show all</a>	<p>1. In the past two decades, our understanding of the human genome has increased at remarkable rate. This expansion of our genetic knowledge promised to resolve longstanding questions regarding the origins and evolution of mankind (Eller, 2002). <b>However</b>, after nearly fifteen years of research, our understanding remains limited; genes have raised far more questions than they have answered. Two models, proposed nearly thirty years ago, still hold weight in the anthropological community: the Multiregional and the Replacement model of the evolution of mankind (Eller, 2002). A critical review of various studies concerning the two models and, specifically, of recent genetic evidence concerning the interaction of <i>Homo sapiens</i> and <i>Homo neanderthalensis</i> (Caramelli et al., 2003; Gibbons, 2001; Kahn &amp; Gibbons, 1997; Krings et al., 2000; Krings et al., 1997; Ovchinnikov et al., 2000), will reveal that - while both models are far too simplistic - a pattern closer to Replacement is more likely.</p>		
<a href="#">BIO.G0.05.1</a>	Mn (III) TPPS4: A Metallophorphyrin Used for Tumor Identification in MRI	Biology	Report
	<p>1. concentration. It follows that Magnetic Resonance Imaging is essentially a measure of proton density. When the magnetic field varies linearly across an object, the signal intensity is proportional to proton concentration. Tumor identification is possible because of differences in relaxation rates of bulk water of normal tissues and tumors. <b>However</b>, the difference is usually no more than 2-3 fold. T<sub>1</sub>, longitudinal relaxation time, is the taken by a proton to align with the external magnetic field. As referenced by Runge and coworkers, longitudinal relaxation rate (1/T<sub>1</sub>) can be described by the following equation:</p>		
<a href="#">BIO.G0.06.1</a>	Global Reproductive Strategies of Tursiops and Stenella (Family Delphinidae)	Biology	Report
1 of 11 hits <a href="#">Show all</a>	<p>1. Members of the family delphinidae, which includes dolphins, killer whales, pilot whales, and many other species, are viviparous mammals. Therefore, it is expected that females will generally be choosy, males will be competitive, and the mating systems will probably be either promiscuous or polygynous. Delphinids are usually sexually dimorphic-males are generally larger than the females and may have a different shape (Animal Diversity Web, 1999 and Tolley et al. 1995). In general, males may be larger than females if male-male competition is an important selecting factor in reproduction. If males combat for access to the females, larger males will have an advantage, and be more likely to pass their genes that code for large size on to the next generation. <b>However</b>, because delphinidae encompasses so many genera and has a worldwide distribution (Animal Diversity Web, 1999), different mating strategies exist and male-male competition is not the sole determining factor for a successful mating for <i>Stenella</i> and <i>Tursiops</i>. The mating systems of the genus <i>Tursiops</i> (which includes bottlenose dolphins) and the genus <i>Stenella</i> (which includes spinner dolphins) have been studied more than many other genera, due to their tolerance of people and relatively easier access. Therefore, the focus of this review will be only on those two genera.</p>		



## **In sum, these are some ways teachers can include corpus in their language classes:**

---

- Developing materials based on language patterns identified in corpus studies.
- Adding corpus examples to exercises or using corpora to inform the development of class materials.
- Helping students explore different corpora in the classroom.

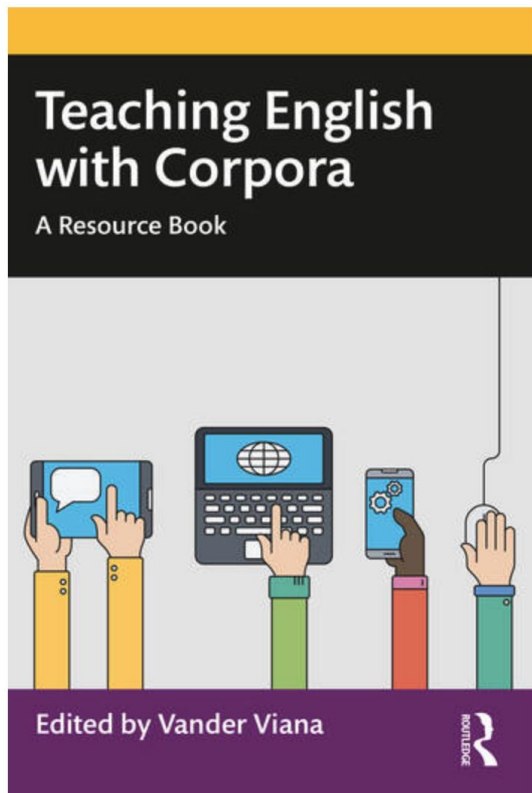
# Revisiting the advantages of using Corpora in the classroom

---

- Authenticity: Corpora allows students to explore language in context
- Autonomy: Students can answer their own language questions
- Specificity: Teachers can use corpora that reflect the type of discourse or disciplinary practices that students want to produce

# Off the shelf corpus activities

<https://pressbooks.pub/elenlefol/>



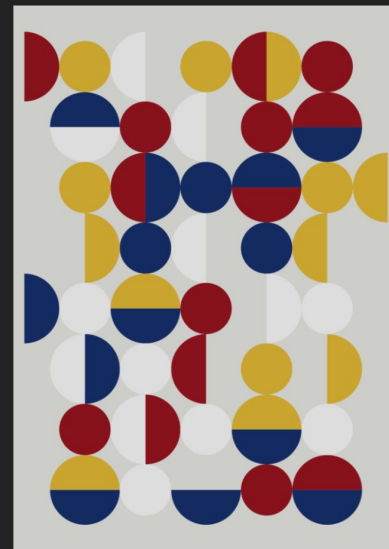
## Creating Corpus-Informed Materials for the English as a Foreign Language Classroom

A step-by-step guide for (trainee) teachers using online resources

Elen Le Foll (Ed.)

 Creative Commons Attribution NonCommercial

[READ BOOK](#)



# Off the shelf corpus activities

<https://uq.pressbooks.pub/using-language-data/>

## Using Language Data to Learn About Language: A Teachers' Guide to Classroom Corpus Use

Paula Tavares Pinto; Peter Crosthwaite; Carolina Tavares de Carvalho; Franciele Spinelli; Talita Serpa; William Garcia; and Adriane Orenha Ottaiano

Using language data to learn about language: A teachers' guide to classroom corpus use contains English, Portuguese and Spanish teaching resources to be used in international language classes. These resources can be used in classrooms with internet access (hands on) or without (hands off). This collection of ready-made DDL lesson plans is curated to help both new and experienced



### USING LANGUAGE DATA TO LEARN ABOUT LANGUAGE

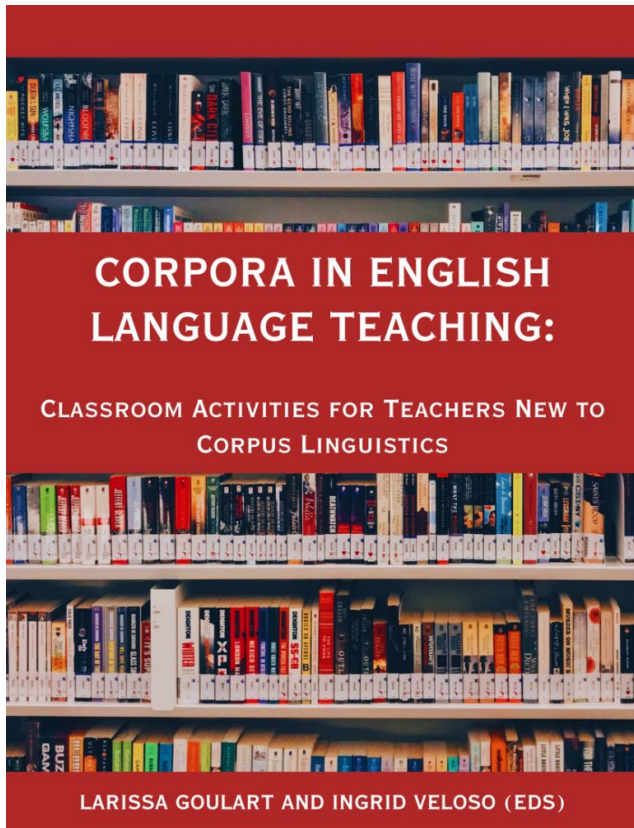
A Teachers' Guide to Classroom Corpus Use



[Download this book](#)



# Off the shelf corpus activities



## Corpora in English Language Teaching

Classroom Activities for Teachers New to Corpus  
Linguistics

Larissa Goulart and Ingrid Veloso



# Corpora in English Language Teaching

## I. Tutorials

### 1. Sketch Engine Tutorial (using the British National Corpus)

Ingrid Veloso

### 2. Corpus of Contemporary American English (COCA) Tutorial

Ingrid Veloso

CLASSROOM ACTIVITIES FOR TEACHERS NEW TO  
CORPUS LINGUISTICS



LARISSA GOULART AND INGRID VELOSO (EDS)

Larissa Goulart and Ingrid Veloso



# Corpora in English Language Teaching

## I. Tutorials

### 1. Sketch Engine Tutorial (using the British National Corpus)

Ingrid Veloso

### 2. Corpus of Contemporary American English (COCA) Tutorial

Ingrid Veloso

## II. Tasks

### 3. Using gender-related vocabulary

Lara Batista

### 4. Using COCA to explore the differences between 'through' and 'toward'

Bruna Andrade; Gabriela Britto; Rita De Leeuw; and Raquel Silva

### 5. How to use 'in', 'on', and 'at' when talking about time

Tess Coffey; Júlia Lisboa; Maria Moysés; and Mariana Trindade

### 6. What can we learn about "take", "have", and "make" from COCA?

Julia Ribeiro; Johanna Tumux; Marcelle Pesenti; and Renata Zimny

### 7. "You didn't cheat, but you're still a traitor": Using COCA to explore the difference between synonyms

CLASSROOM ACTIVITIES FOR TEACHERS N  
CORPUS LINGUISTICS



LARISSA GOULART AND INGRID VELOSO (E

# Off the shelf corpus activities

---

## Corpus for Schools

<http://wp.lancs.ac.uk/corpusforschools/>

This is a project from Lancaster University that brings together corpus linguistics, applied linguistics, teachers and material writers to develop teaching materials and online platforms that incorporate corpus-based findings as well as direct access to corpora to teach about how English is used in real life situations. The materials were developed both for A-level English Language classes as well as for teaching English as a foreign/second language classes.

## CROW Resources and Activities for Teachers

<https://writecrow.org/crow-for-teachers/pedagogical-materials/>

A database of corpus-based activities created by the CROW team.

## Creating Corpus-Informed Materials for the English as a Foreign Language Classroom

<https://pressbooks.pub/elenefall/>

This book is a bank of ideas and resources to inspire teachers to use corpus tools in the language classroom.

## BAWE Quicklinks Project

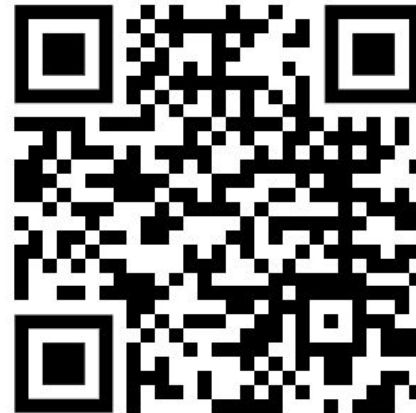
<https://bawequicklinks.coventry.domains/>

This site is designed as an aid for teachers who'd like to introduce students to concordances (see example below) that can help raise their awareness of how English works.

## Kibbitzers

<https://lexically.net/TimJohns/>

These are ideas Tim created for language teaching, usually based on problems he had met in his work in English for Academic Purposes at Birmingham University.



<https://www.coralcorpuslab.com/corpus-and-teaching>





To join the **Corpus Revolution!**



**Thank you!**

**Larissa Goulart**  
*Montclair State University*  
goulartl@montclair.edu



# Corpus Linguistics in Teaching Materials

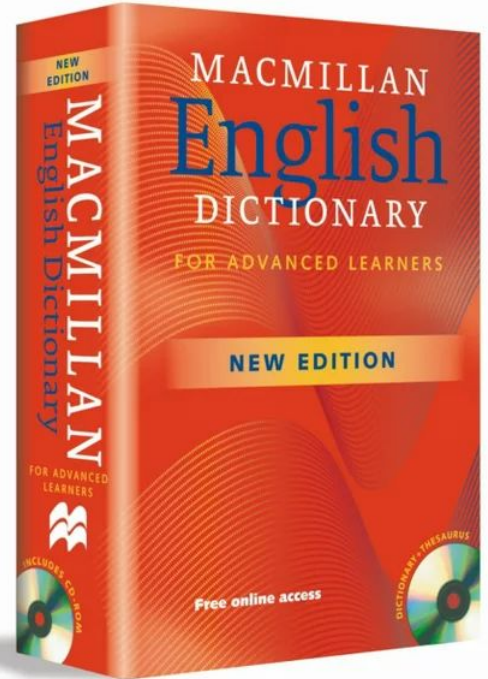
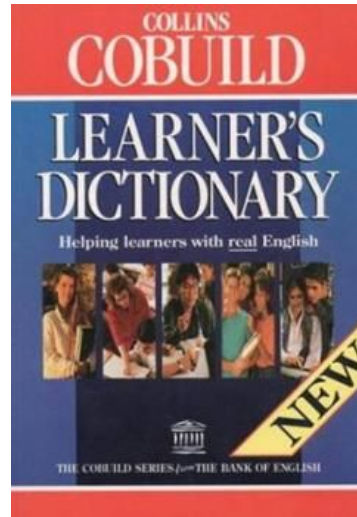
---

- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

---

## - Dictionaries



# Corpus Linguistics in Teaching Materials

---

- Dictionaries
- **Wordlists**
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

## - Wordlists

### A New Academic Vocabulary List FREE

Dee Gardner ✉, Mark Davies

*Applied Linguistics*, Volume 35, Issue 3, July 2014, Pages 305–327,  
<https://doi.org/10.1093/applin/amt015>

From Yang, M-N. (2015). A nursing academic word list. *English for Specific Purposes*, 37, 27-38.

#### NURSING ACADEMIC WORD LIST

analyze	aware	anatomy	component	comprehensive
assess	adapt	admission	cue	correspond
area	aggression	advocate	communicate	considerable
affect	adverse	alleviate	contribute	colleague
adolescent	abdomen	appreciate	consume	chemotherapy
approach	assign	analogue	contact	constipation
appropriate	assume	ascertain	complex	cite
available	attach	attenuate	cognition	chart
assist	aid	benefit	consequent	compress
adjust	apparent	baseline	context	compute
adult	autonomy	barrier	concept	covariate

“ Cite Permissions Share ▾

ir new Academic Vocabulary List (AVL),  
llion-word academic subcorpus of the  
pus of Contemporary American English  
to first explore reasons why a new

## The Academic Word List

The Academic Word List is a useful English resource for lecturers and students.

[Averil Coxhead](#) from the [School of Linguistics and Applied Language Studies](#) developed and evaluated *The Academic Word List* (AWL) for her MA thesis. This list is a very useful resource for English for Academic Purposes teachers and learners.

- [AWL Information](#)
- [AWL Sublist Families](#)
- [AWL Headwords](#)
- [AWL Most Frequent Words in Sublists](#)
- [AWL Tests](#)
- [Useful links](#)

If you have questions about the AWL, contact



### [Prof Averil Coxhead](#)

Professor · Applied Linguistics and TESOL  
School of Linguistics and Applied Language Studies

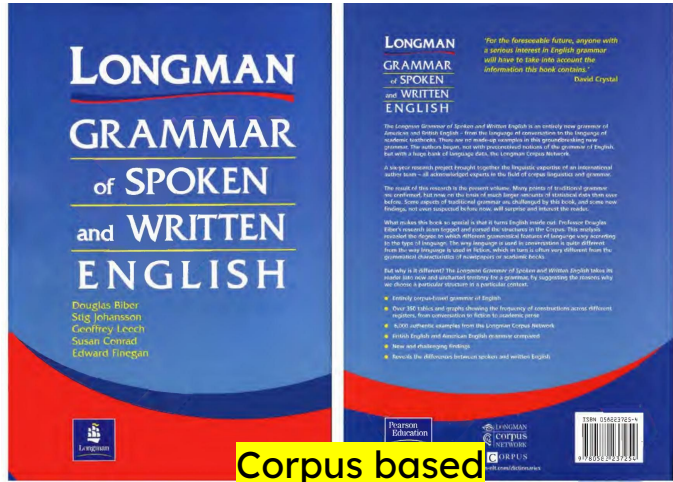
# Corpus Linguistics in Teaching Materials

---

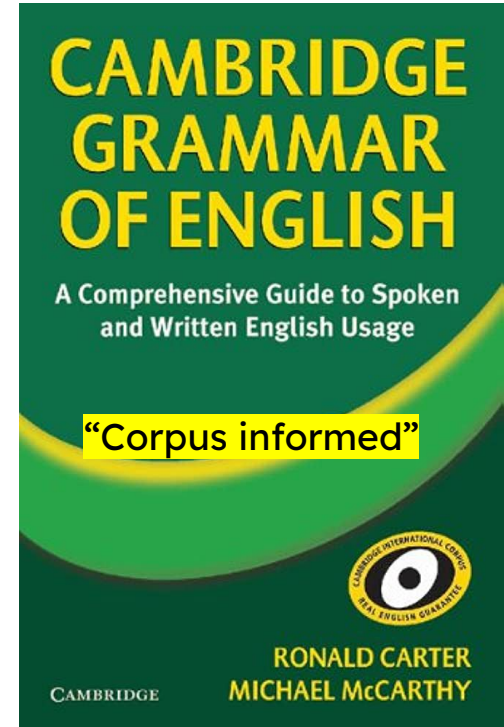
- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

## - Reference Grammars



Corpus based





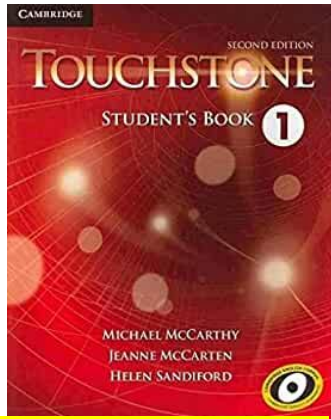
# Corpus Linguistics in Teaching Materials

---

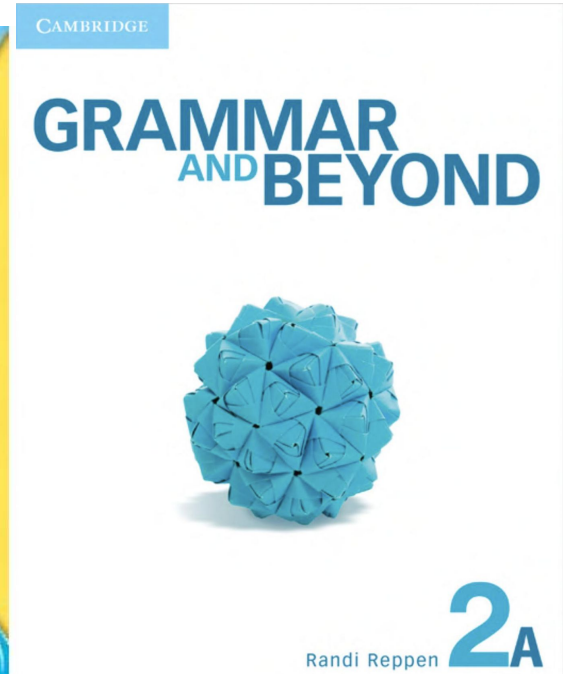
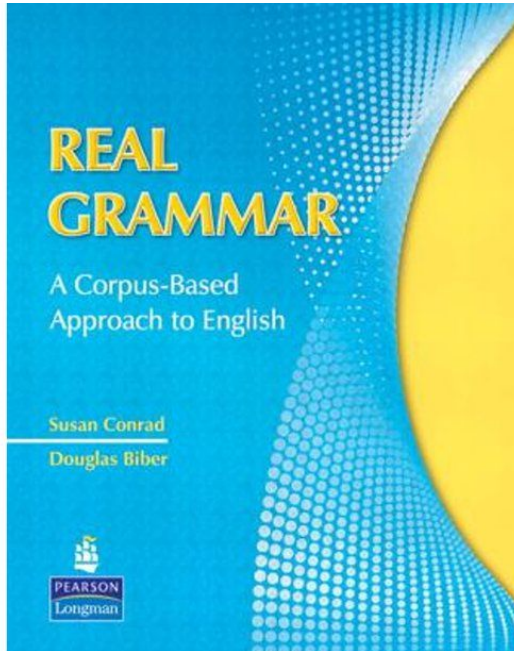
- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

## - Textbooks



“Corpus informed”



# Corpus Linguistics in Teaching Materials

---

- Dictionaries
- Wordlists
- Reference Grammars
- Textbooks
- Assessment

# Corpus Linguistics in Teaching Materials

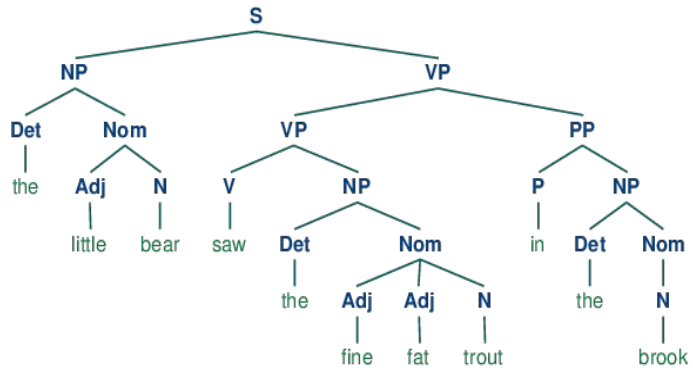
---

## - Assessment

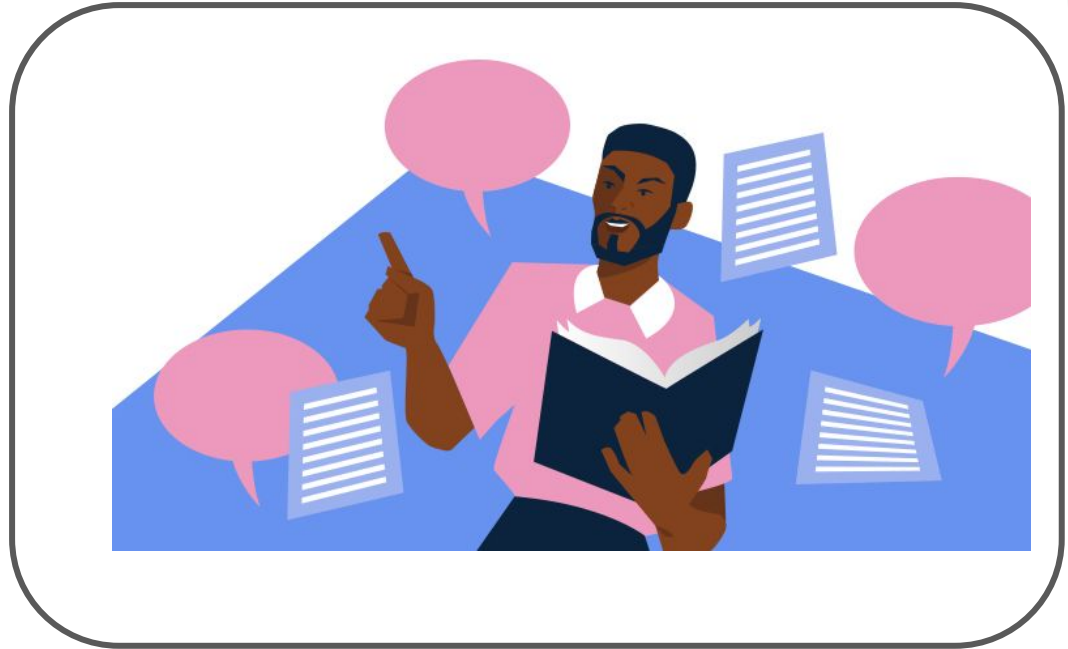
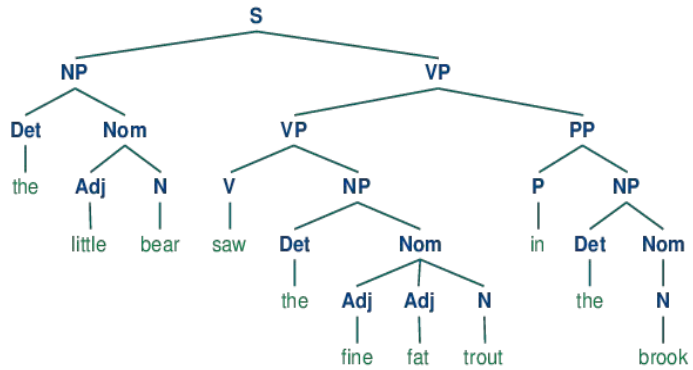
Using Corpora to:

- Develop items
- Identify learner's challenges
- Evaluate validity of the exam

# Language Studies



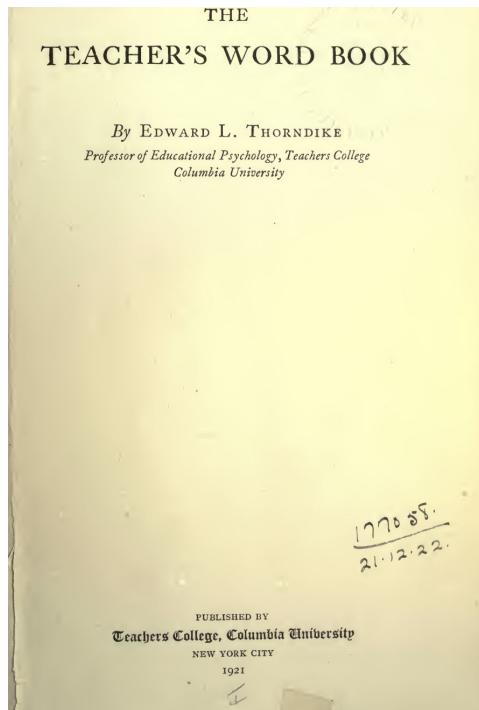
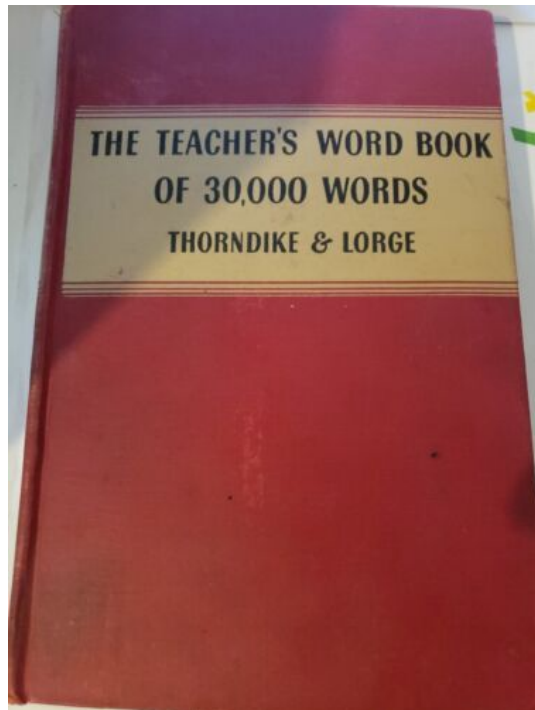
# Language Studies



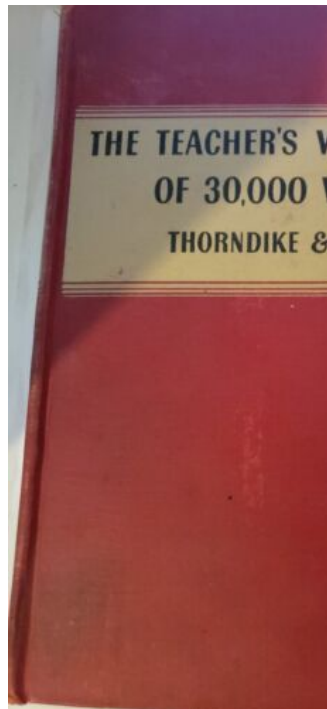
Language in use

# Corpus in the classroom: Before electronic corpora

---



# Corpus in the classroom: Before electronic corpora



feud	6		financial	11	5b
fever	33	2b	find	131	1a2
few	99	1a3	fine	108	1a3
fib	3		fineness	5	
fiber	13	5a	finger	54	1b
fickle	8		finish	72	1b
fiction	9		fir	16	4a
fiddle	9		fire	111	1a3
fideliity	5		firearms	5	
Fido	4		firebrand	6	
fie	5		firefly	4	
field	92	1a4	fireman	7	
fiend	12	5a	fireplace	10	5b
fierce	41	2a	fireproof	4	
fierceness	5		fireside	7	
fiery	21	3b	firm	53	1b
fife	8		firmament	10	5b
fifteen	43	2a	firmness	5	
fifteenth	10	5b	first	136	1a2
fifth	42	2a	first-born	8	
fiftieth	4		fish	66	1b
fifty	46	2a	fisher	10	5b
fig	15	4b	fisherman	15	4b
fight	69	1b	fishhook	3	
fighter	8		fist	20	3b
figure	68	1b	fit	68	1b
file	26	3a	fitness	6	
filer	3		five	101	1a3
filial	6		fix	55	1b
fill	94	1a4	fixture	7	
filler	5		flag	43	2a
fillet	5		flake	13	5a
film	10	5b	flame	36	2a
filth	7		Flanders	4	
filthiness	3		flange	4	
filthy	10	5b	flank	10	5b
fm	7		flannel	9	
final	40	2a	flannelette	4	
finally	40	2a	flap	13	5a
finance	7		flare	5	

[42]



Most frequent in words  
in English based on a  
corpus of books and  
legal documents.

The list was compiled  
manually.