

## Using Corpus-Based Web Resources – How Can They Be Manipulated For Research in Language Patterns

Le Thi Giao Chi<sup>1</sup>, Nguyen Bich Dieu<sup>2</sup>

<sup>1</sup>*Faculty of Foreign Language Teacher Education, University of Foreign Language Studies -  
the University of Danang, 550000 Danang, Vietnam*  
Tel: (+84) 236 3699 177 / (+84) 905 157 325

Email: [ltgchi@ufl.udn.vn](mailto:ltgchi@ufl.udn.vn)

<sup>2</sup>*Faculty of Foreign Language Teacher Education, University of Foreign Language Studies -  
the University of Danang, 550000 Danang, Vietnam*  
Tel: (+84) 236 3699 177 / (+84) 934 863 568

Email: [nbdieu@ufl.udn.vn](mailto:nbdieu@ufl.udn.vn)

---

**Abstract:** *The introduction of Internet and the emergence of available web resources have facilitated endless efforts by language teachers to do research in language patterns. Web resources in general and corpus tools in particular have enabled large samples of language to be explored for better insights into the nature of language in use in all its forms and its uses. While corpora are normally assumed to be in the hands of lexicographers whose job is to inform dictionaries or grammar books, arguments may have arisen around why such end-users as language teachers and learners cannot make use of these innovative tools. This paper adds to this on-going debate by discussing approaches to using corpora as a reference point for language teaching and research. It shows potentials of manipulating common data-driven web tools for research in language patterns. It also explores some pathways for language teachers to examine aspects of language in use through authentic texts accessed via corpus tools. Results revealed from corpus search and generalizations made from manipulated data, and in this case noun and verb patterns and grammatical metaphor, in COCA and BNC can further showcase the inexhaustible implications of using web resources for enhanced language teaching and research.*

**Key words:** *web resource; corpus tools; corpus data; language patterns; grammatical metaphor.*

---

### I. Introduction

The use of educational technology has opened up multiple pathways for innovations in language teaching and learning. Many a language teacher and education practitioners have counted on technologies and resources available online to constantly update their teaching content and enhance their instructional experience. In the arena of language research, web resources and tools have enabled large samples of language to be explored for better insights into the nature of language in use.

There has been quite a good body of literature done in the usefulness of Internet resources and tools for enhanced language teaching and learning (e.g. Kamba 2007; Son, 2011, Pim, 2013; Geladze 2015; Ahmadi, 2018). Kamba (2007), for example, examined the usefulness of the Internet as a tool for interactive learning, teaching and research. Son (2011) explored a number of online tools that can be used in second or foreign language teaching and learning. Pim (2013) recognized the prevalence of Information and Communications Technologies (ICTs) in supporting English language lessons, helping teachers to adopt innovative approaches in order to “meet the needs of the young technocrats growing up within an increasingly globalised world” (p.17). Geladze (2015) recognized the important role of modern informational technology and resources in innovating the teaching and learning process and in upgrading teaching on the scientific-methodological level. Erenchinova and Proudchenk (2017) acknowledged the use of Web resources in significantly improving the efficiency of learning a foreign language, enabling students to form and develop linguistic and communicative skills. Ahmadi (2018) further exhibited the impact of the Internet in advancing research, teaching and learning abilities as well as techniques that can be utilized by language teachers and researchers.

While it is important that the main goal of teacher as researchers is to search innovative ways for improving teaching and learning, knowledge on language patterns driven by large data size can really inform language teaching, thus adding to enhanced language learning. Without doubt, this depends on the new techniques and/or specific pedagogic technologies which can be realized by means of computer-assisted ways (Braun et al, 2006). This by all means brings to the fore the use of corpus technology, which is meant to inform language learning

and teaching with corpus-based observations revealing patterns of real language use, thus helping to uncover several problems of language in authentic operations. This paper, thus, adds to the on-going discussions of the multiple uses of internet resources, or rather, corpus-based web resources with case studies revealed for language teachers as researcher to explore aspects of language in use through authentic texts available online and/or accessed via corpus tools for research in language pedagogy.

## II. Corpus Linguistics and Corpus-based Pedagogies – An Overview

Corpus linguistics is an increasingly popular field of linguistics, and according to Baker (2010, p. 93), it refers to “the analysis of very large collections of electronically stored texts, aided by computer software”. In the position of McEnery and Hardie (2012), corpus linguistics deals with “some set of machine-readable texts which is deemed an appropriate basis to study a specific set of research questions” (2012, p. 1). Corpus linguistics is therefore a methodology, or an approach used to investigate linguistic phenomena rather than a sub field within linguistics (Ngula, 2018).

As its name implies, a corpus is a large ‘body’ of texts stored electronically. A general corpus is one that includes a variety of text types, ranging from written texts, spoken texts, or both, and very often it represents a national, regional or sub variety of a language. There are several general corpora of approximately a million words, such as the Lancaster-Oslo-Bergen (LOB) written corpus, and others of a much bigger size that include both written and spoken texts, such as the 100 million-word British National Corpus (BNC) and the over 450 million-word Contemporary Corpus of American English (COCA).

In fact, there has been a great consensus on the increasing availability of corpora and analysis tools that has made it easier to manipulate corpora in much wider ways, many of which are seen as relevant

not only to language research but also to language pedagogy. With increasing interest in corpus technology for pedagogical uses, efforts have now been seen in integrating existing corpora, corpus methods and tools into teaching practice, and so have methods of exploring and expanding existing web tools geared for pedagogical research questions and goals. Thus, it can be said that corpus linguistics could revolutionize language teaching, by fundamentally changing the ways we approach all areas of pedagogy, including materials development, curriculum design, and teaching methodology.

## III. Corpus-Based Web Tools For Language Research – Cases With COCA And BNC

All the functions of BNC and COCA can be useful in exploring language patterns. Why **SEARCH** is essential for getting to know the impressions of how language works, other functions such as **KWIC**, **COLLOCATES**, and **COMPARE** would be great to enable language teachers and students to study grammatical patterns, especially collocations, and to be able to explain why such patterns occur the way they are, and to get exposed to a great source of English grammar patterns as they are used in authentic contexts in different genres, spoken and written, academic and non-academic, fiction and non-fiction. Figure 2 shows a snapshot of some corpus functions.



Created by Mark Davies, BYU (Google Scholar): [Overview](#), [search types](#), [looking at variation](#), [corpus based resources](#).

The most widely used online corpora – more than 130,000 distinct researchers, teachers, and students each month:

English	# words	language/dialect	time period	compare
Web: The Intelligent Web-based Corpus	1.6 billion	US/CA/UK/EU/JA/IN/Z	2017	<a href="#">Info</a> (中英)
News on the Web (NOW)	7.47 billion+	33 countries / Web	2010-present	
Global Web-based English (GloWBE)	1.3 billion	33 countries / Web	2012-15	
Wikipedia Corpus	1.5 billion	English	2014	<a href="#">Info</a>
Harvard Corpus	1.5 billion	British (parliament)	1803-2005	<a href="#">Info</a>
Early English Books Online	755 million	British	1475-1600	
Corpus of Contemporary American English (COCA)	560 million	American	1990-2017	****
Corpus of Historical American English (COHA)	460 million	American	1810-2009	**
The TV Corpus <b>NEW</b>	325 million	US/CA/UK/EU/JA/IN/Z	1950-2018	<a href="#">Info</a>
The Movie Corpus <b>NEW</b>	200 million	US/CA/UK/EU/JA/IN/Z	1930-2018	<a href="#">Info</a>
Corpus of US Supreme Court Opinions	130 million	American (law)	1790-present	
TIME Magazine Corpus	100 million	American	1923-2006	
Corpus of American Soap Operas	100 million	American	2001-2012	*
British National Corpus (BNC/BNCX)	100 million	British	1980s-1993	**
Swathy Corpus (Canada)	50 million	Canadian	1970s-2000s	
CORE Corpus	50 million	Web registers	2014	

*Figure 1. Introducing English Corpora*

Select a section: (sub-)genre or (set of) year(s). [Click here](#) for more examples of change over time.

*ize verbs in ACADEMIC	Past tense verb + <i>over</i> in SPOKEN
*ment in ACADEMIC	Nouns near <i>green</i> in 2000-2009
Adj + <i>track</i> in NEWSPAPERS	Noun near <i>chair</i> in FIC
Hard + NOUN in MAGAZINES	Synonyms of <i>smart</i> in FICTION
Verbs in MAGAZINES-Sports	Nouns in NEWSPAPERS-Money
Adjectives in ACADEMIC-Medicine	Adverbs in FICTION-Movies

(Optional) Select a second (set of) section(s) against which to compare the sections chosen above

*ize verbs in ACAD vs MAG	Past tense verb + <i>over</i> in FIC vs ACAD
*ment in ACAD vs FIC	Nouns near <i>green</i> in 2005-2015 vs 1990-1999
Adj + <i>track</i> in NEWS vs SPOK	Nouns near <i>chair</i> in ACAD vs FIC
Hard + NOUN in MAG vs ACAD	Synonyms of <i>smart</i> in FIC vs NEWS
Adj in NEWS-Money vs NEWS	Nouns in MAG-Sports vs MAG
Adj in ACAD-Medicine vs ACAD	Phrasal verbs with <i>up</i> 2005-2015 vs 1990-1994

These corpus-based web tools can also be used for preparation of paper-based grammar materials. In fact, language teachers and learners can use BNC and/or COCA, especially with the functions of Collocates, KWIC and Compare to find ample examples of real English language in use with confusing words like ‘*come*’ vs. ‘*go*’, ‘*bring*’ vs. ‘*take*’ or ‘*small*’ vs. ‘*little*’, and so on to design grammar exercises or quizzes for students. **COLLOCATES** can also provide insight into repeated grammar patterns that help teachers explore rules of grammar and that enable teachers to explain these rules to students with confidence and ease once real examples from real contexts are available.

BNC and COCA could also be great tools that facilitate self-regulated and collaborative learning. With a driving question to know about the grammar rules of some particular language phenomena like subjunctive or zero-conditionals or patterns with the pseudo subject ‘it’, etc., teachers can promote project-based learning even in teaching grammar and students are encouraged to do their projects, exploring COCA to find answers to their ‘driving questions’ or an assigned case of grammar, making notes, taking examples, comparing the use of grammar in different registers and prepare a presentation or a report using results from corpus-based tools presented in charts. There are huge applications for teaching grammar and vocabulary but what is important is that students need sufficient scaffolding from teachers to be able to explore the tool to their fullest extent.

#### IV. Corpus-based Web Resources and Research in Language Patterns – Some Case Studies

##### 4.1. A Case Study with SPEND Patterns

In order to prepare for the material development, I did some research into the existing syllabus to find the topic which is considered as relevant to my students in the context of learning integrated language skills. As the communicative language teaching approach is preferable, I found that some practice of grammar in use would be relevant for students to lead in more meaningful communication about the topic. As Plans and Holidays is one of the topics in the course of General English, I noted it down and browsed the internet resources as suggested for this activity. I could see some relevance in the use of grammar to talk about holiday plans and the EVERYDAY GRAMMAR available at Learning English.voanews.com especially with the TWO -ING WORDS IN A ROW. I then visited another resource suggested – American English Facebook (See Figure 3) and found an interesting match with a section on SPECIAL EXPRESSIONS + VERB + -ING with certain verbs like ‘*spend*’, ‘*waste*’, ‘*have a good/bad time*’ to be used with this pattern. As these verbs can be used to describe how people spend their time and to plan a holiday, I decided to make some notes of the instances given on the webpage and design some meaningful practice for my students with gap-fill, sentence completion task before proceeding to some more communicative practice whereby students can make good use of the verb patterns to talk about holiday plans.





Figure 3. Screenshots of special expressions with Verb + ING

To further explore language patterns with verbs denoting holiday plans, I explored COCA and retrieved several authentic instances to adapt into the materials that I could design for the controlled practice of some grammar points before my students can proceed with more productive tasks – discussions about holiday plans. Students can then be given the links to access <http://corpus.byu.edu/coca/> and explore the use of some key words with -ING used in real context. This can facilitate student autonomous learning and thus clear guidance, or instructions are necessary. For example, students can be asked to note down some special patterns from the concordance lines extracted from COCA when they search one of these key words: 'spend', 'waste', 'consider', 'try', etc. Figure 4 shows a snapshot of concordances with *spend*, for example, revealed in COCA, indicating the patterns in which the verb is used and the similar or different contexts it occurs. This enables students to examine the linguistic contexts whereby the word is used while being informed of the particularly distinctive features characteristic of the word.

Corpus of Contemporary American English									
SEARCH			FREQUENCY		CONTEXT			ACCOUNT	
106	2007	MAG	Sport Illustrated	A   B   C	...people are going to be wondering: <b>Why should we</b>	spend	but	that	time
107	2012	SPOK	HBS: PBS Newshour	A   B   C	the entertainment spending is a giant wave. And <b>we</b>	perpetually	spend	but	time
108	2007	ACAD	ForeignAffairs	A   B   C	somehow, corn oil, and corn <b>spend</b>	but	time	on	the
109	1992	ACAD	Education	A   B   C	side for every 4 to 6 teachers. 1,846 4th <b>Teachers</b>	spend	but	time	on
110	2006	SPOK	NBC Today	A   B   C	time, but price it out for <b>how much you should</b>	spend	but	time	on
111	1998	MAG	USCatholic	A   B   C	to dig the shot out of his <b>bars</b>	he	would	spend	but
112	2010	ACAD	SocialWork	A   B   C	for close to four years, in Uganda, <b>children</b>	would	spend	but	time
113	1998	MAG	Forbes	A   B   C	disparities, I assumed. But I <b>was</b>	surprised	to	spend	but
114	2006	NEWS	NYTimes	A   B   C	stealing food. It was not until her <b>son</b>	came	to	spend	but
115	1996	MAG	Esquire	A   B   C	show, Jackie talks about the English <b>royals</b>	and	how	they	spend
116	1990	NEWS	WashPost	A   B   C	offseason and last. "So now <b>you</b>	are	forced	to	spend
117	2013	SPOK	ABC 20/20	A   B   C	don't know what would be going <b>through</b>	his	mind	to	spend
118	2013	FIC	BkZeldaFitzgerald	A   B   C	would as soon dance naked in front <b>of</b>	the	courtroom	as	spend
119	1990	SPOK	ABC Nightline	A   B   C	you take that chance. And they <b>really</b>	do	it	even	spend
120	2004	NEWS	Atlanta	A   B   C	him after a birthday party, but <b>she</b>	was	scheduled	to	spend
121	1996	MAG	Skiing	A   B   C	operates in packaged weekly cycles, where all the <b>guests</b>	spend	but	time	on
122	2017	FIC	BkManderMoriMalicious	A   B   C	to bridge home through last night's <b>bursting</b>	of	love	to	spend
123	1994	MAG	Horiculture	A   B   C	in wintertime folktrams, clivia come to the <b>rescue</b>	to	spend	but	time
124	2004	FIC	Archer	A   B   C	on in the dean channel. These <b>same</b>	flowers	also	spend	but

Figure 4. Concordances of SPEND in COCA

#### 4.2. A Case Study with TRY patterns

In order to identify the patterns of language of **try** as its part of speech varies, the *List* and *KWIC* functions in COCA are exploited to compare the frequency of the word **try** and observe their concordance lines. With the *List* function, I put the word **try** in the SEARCH box. By clicking FIND MATCHING STRINGS, the frequency comes up to 138 333. By clicking the word **try**, it reveals the authentic examples with **try** in the concordance lines and the real contexts in which the word occurs. When tagging POS to the word with verb.all – try.[v\*], and with noun.all – [try.[nn\*], **133 736** and **4589** occurrences come up on the screen respectively, which shows that **try** is more often used as a verb than as a noun.

With this initial finding, I continued to explore the word **try** used in real contexts, and this throws some very interesting insights into the common patterns or phrases whereby the word is used. When used as a verb, **try** can function both as an intransitive verb or a transitive verb. Being an intransitive verb, **try** is found predominantly with the adverb **again** as in '**Try again**', '**Please try again**' '**Why don't you try again**' or **out** as in '**OK. Come try out**' or '**try out again**' '**try out over the weekend/underwater/for the next season**'. When acting as a transitive verb, **try** needs an object be it a noun, a noun phrase (e.g. **try innovative ideas**), or a gerund. Interestingly, I found out that the phrase **try out** has been used quite often. As I searched the entire phrase, I could find **1207** occurrences of **try out** in the corpus with plenty of interesting examples with concordance lines - **try out** + N/NP (e.g. **try out the position / new powers / different ideas / the new fundraising platform**; or **try out someone that knows how to really find the love in any situation**, etc.) or **try out** + V-ING like the following

'If only the youthful years were set aside for the children to **try out** being a boy one day, a girl the next, an angel the day after, a monk or an astronaut on the following days' (Context from COCA)

When used as a noun, **try** is often preceded by a determiner (e.g. **a try**) occurs in some common patterns such as

- give it a **try**
- worth a **try**
- make a **try**
- give one more **try**
- attributive adj + **try** (a real **try** / nice **try** / good **try** / his best **try** / his latest **try**)
- on his first **try** / on the second **try**

The KWIC function with the two most common patterns with **try** where the frequencies can reveal which of the two is more prevalent in the corpus. While **worth a try** occurs up to 256 times, its counterpart **give it a try** is even much more commonly used in the corpus, especially in spoken discourses (619 occurrences). Also, the results from KWIC function of COCA show that **worth a try** often occurs with the verb 'be' with or without adverbs showing certainty like *certainly* or *definitely*.

It is	worth a <b>try</b> .
It was	
It may be	
It might be	
It's certainly/definitely	

While patterns **It + BE + worth a try** are used predominantly in fiction, patterns with **give it a try** are then found in more various registers: fiction, news, spoken, and magazine. When being used in spoken domain, **give it a try** often comes with *Please / I'll / Why don't you* [give it a try] to show ways of giving advice or making suggestions.

SEARCH

FREQUENCY

CONTEXT

CONTEXT +

Source information:

Date	1996 (May)
Publication information	Vol. 173, Iss. 3; pg. 96, 24 pgs
Title	Solving steering-pull problems
Author	Wessler, Paul
Source	MAAG: Popular Mechanics

Expanded context:

drag, if you get much more drag at one front wheel than the other, you have a brake-system problem -- either a bad caliper or something wrong with the hydraulic system such as a crimped tube or even bad master cylinder. Equal or no drag on both from tires? Switch them left to right. If the pull is in the opposite direction, get new tires. Even if the pull changes somewhat, but not necessarily all the way in the opposite direction, the tires will are the most probable cause of the pull, and a new pair might be **worth a try** if nothing else pans out. If you're lucky, however, the pull will go away and you can pat yourself on the back for a cheap fix. Stop here if you're checking for a cause of torque steer. Let's try one more thing for a tire-related pull. If there's no change after the front-tire swap, next switch the wheel on the driver's-side front with the one on the driver's-side rear. If that makes a significant change, get a

### 4.3. A Case Study with Confusing Words – *LITTLE* vs *SMALL*

With the display of COCA, collocates with **little** and **small** are shown in two lists: W1-LITTLE vs. W2-SMALL and V1-SMALL vs V2-LITTLE. This enables the noting down of most common collocates with either W1 or W2 or with both. The word BIT, for example, is found in collocates with both **little** and **small**. **Little in a little bit** is found with a frequency of 192.9 or 40 889 occurrences (*feel a **little** bit of annoyance/ a **little** bit arrogant/tired/crazy/...*). *Bit* is also found with **small**, though with lower frequency – 105.1 or 212 occurrences. The word BUSINESSES is also recognized to go with both **little** and **small**, though of course, with the meaning of ‘limited in the amount of activity’ it collocates with **small** up to 3365 times. The phrases *small businesses* are common in today business world with 3741 occurrences (See Figure 7). Similar phrases of this meaning are *small producers / employers/farmers/donors, etc.* *Small* can also be found in collocates with *small sums, small scopes, small tumors* whereby its central meaning is ‘small in size’.



Corpus of Contemporary American English					CONTEXT				
SEARCH	FREQUENCY				CONTEXT	CONTEXT +			
43. MATTERED	104	0	208.0	113.3	43. STATISTICALLY	53	0	110.0	201.9
44. RELAXED	104	0	208.0	113.3	44. INFLUENTIAL	53	0	106.0	194.6
45. STRAIGHTER	103	0	206.0	112.2	45. MAGELLANIC	52	0	104.0	190.9
46. CAESARS	102	0	204.0	111.1	46. SPATULA	52	0	104.0	190.9
47. WASTED	100	0	200.0	109.0	47. SELECT	51	0	102.0	187.2
48. SAIGON	99	0	198.0	107.9	48. FAMILY-OWNED	50	0	100.0	183.6
49. CONFUSING	98	0	196.0	106.8	49. SCOPES	50	0	100.0	183.6
50. BIT	4088	212	192.8	105.1	50. INFINITESIMALLY	49	0	98.0	179.9
51. OVERSIGHT	96	0	192.0	104.6	51. RESIDENTIAL	48	0	96.0	176.2
52. RIDICULOUS	96	0	192.0	104.6	52. FRAMED	47	0	94.0	172.5
53. CHILLY	94	0	188.0	102.4	53. MID-SIZED	47	0	94.0	172.5
54. DOZZY	94	0	188.0	102.4	54. SUMS	47	0	94.0	172.5
55. SLEPT	94	0	188.0	102.4	55. THIRLY	47	0	94.0	172.5
56. ANNOYED	92	0	184.0	100.2	56. BUSINESSES	3741	40	93.5	171.7
57. QUICKER	92	0	184.0	100.2	57. ADJACENT	46	0	92.0	168.9
58. EXPLANATION	91	0	182.0	99.2	58. EARTHQUAKES	46	0	92.0	168.9
59. FRIGHTENING	89	0	178.0	97.0	59. PURCHASES	46	0	92.0	168.9
60. MERRY	86	0	172.0	93.7	60. TUMORS	46	0	92.0	168.9

Figure 7. Collocates with 'LITTLE' vs 'SMALL' in COCA

This COMPARE function thus helps us explore words that collocates with one rather than the other word. With the distinctive function as an adverb that precedes an adjective or a comparative (*a little guilty/younger/louder/faster/slower, ...*), **small** doesn't reveal any matches. However, in cases where the meaning conveyed is 'limited in size/amount/scope or smaller than the average size' we can see more collocates with **small** like small budgets (27)/amounts (1126)/quantities (283)/percentages (1231)/proportions (290), .... Interestingly, SAIGON, the name of the biggest city in Vietnam occurs 97 times in **Little Saigon**, which may indicate reference of a small Vietnamese community in the US. Similarly, the proper name **Little League** occurs up to 1445 times in the News and Spoken texts of COCA with *Little League Game* or *Little League Coach*.

British National Corpus (BNC)					CONTEXT				
SEARCH	FREQUENCY				CONTEXT	CONTEXT +			
32. LUSH	5	0	10.0	30.1	32. COMPETENCE	18	0	30.0	10.0
33. NOMINAL	5	0	10.0	30.1	33. PATTERNS	18	0	30.0	10.0
34. NON-ELECTED	5	0	10.0	30.1	34. STYLE	20	1	28.0	9.9
35. ROOTS	5	0	10.0	30.1	35. ASPECTS	14	0	28.0	9.9
36. SLOWING	5	0	10.0	30.1	36. BEHAVIOUR	13	0	28.0	9.8
37. SPRING	5	0	10.0	30.1	37. DETAILED	13	0	28.0	9.8
38. WAGES	16	2	8.0	24.1	38. FIELDS	14	0	28.0	9.9
39. JOBS	8	1	8.0	24.1	39. IMPROVED	14	0	28.0	9.9
40. DENSE	4	0	8.0	24.1	40. LOOKING	13	0	28.0	9.8
41. EXPONENTIAL	4	0	8.0	24.1	41. PARTNERSHIP	13	0	28.0	9.8
42. EXPORT	4	0	8.0	24.1	42. SEXUAL	14	0	28.0	9.9
43. FORECAST	4	0	8.0	24.1	43. SUSTAINABLE	14	0	28.0	9.9
44. FUNGI	4	0	8.0	24.1	44. VOCATIONAL	13	0	28.0	9.8
45. GOVERNMENTS	4	0	8.0	24.1	45. ACTIVITIES	27	1	27.0	9.0
46. GROSS	4	0	8.0	24.1	46. TEACHING	27	1	27.0	9.0
47. HAIR	4	0	8.0	24.1	47. APPLICATION	19	0	26.0	8.6
48. KEEP	4	0	8.0	24.1	48. MODE	12	0	26.0	8.6
49. PROFITS	4	0	8.0	24.1	49. PROCEDURES	13	0	26.0	8.6
50. REGULATION	4	0	8.0	24.1	50. TECHNIQUE	19	0	26.0	8.6
51. RENTED	4	0	8.0	24.1	51. TESTS	19	0	26.0	8.6
52. SELF-EMPLOYMENT	4	0	8.0	24.1	52. TREATMENT	12	0	26.0	8.6
53. SIMULTANEOUS	4	0	8.0	24.1	53. TECHNIQUES	31	2	25.8	8.5
54. STURBLE	4	0	8.0	24.1	54. COMMUNICATION	25	1	25.0	8.3
55. SULPHATE	4	0	8.0	24.1	55. SPECIAL	26	1	25.0	8.3
56. TRANSNATIONAL	4	0	8.0	24.1	56. CHILDREN	72	2	24.0	8.0
57. DEMAND	29	4	7.3	21.8	57. AIRCRAFT	12	0	24.0	8.0
58. BASAL	7	1	7.0	21.1	58. AVAILABLE	12	0	24.0	8.0
59. AGGREGATE	7	1	7.0	21.1	59. COMPUTING	12	0	24.0	8.0
60. INCOMES	7	1	7.0	21.1	60. EVALUATION	12	0	24.0	8.0
61. YAK	7	1	7.0	21.1	61. EVOLUTIONARY	12	0	24.0	8.0
62. HUNDRED	6	1	6.0	18.1	62. FOCUS	12	0	24.0	8.0
63. INFLATION	6	1	6.0	18.1	63. METHOD	12	0	24.0	8.0
64. LIMITS	6	1	6.0	18.1	64. ORGANIZATION	12	0	24.0	8.0
65. NARROW	6	1	6.0	18.1	65. POLITICS	12	0	24.0	8.0

Figure 8. Collocates with 'LITTLE' vs 'SMALL' in BNC

Some interpretations into the use of **little** versus **small** in its collocation can be gained here. While the left section shows chances that certain words can go together with **little** versus **small** with corresponding occurrences and percentages, the right section shows the opposite. It is thus interesting to find that we can see **small** in *small politics*, *small organisations*, *small treatment*, *small test*, *small techniques*, *small procedures*, etc. whereby the number zero in the column denotes that there is no such case of collocates with **little**.

#### 4.4. A Case Study with Nominalisations

With the SEARCH function, both COCA and BNC can reveal results regarding frequencies, context, and also collocations of nominalisations as grammatical metaphor in authentic contexts. With the trial in the search box for such nominalisations ending with -MENT, results extracted from the COCA tool can throw up

most common collocations with the search word being identified with supporting frequencies ranked in the descending order. While such nouns as ‘department’, ‘moment’, ‘element’, ‘segment’ are clearly not the case, nominalisations like ‘development’, ‘treatment’, ‘movement’, ‘assessment’, ‘management’, are at the top of the list with frequencies being shown in Figures 9-10.



Figures 9-10. Frequencies of nominalisations with -MENT in COCA

The CONTEXT function in COCA/BNC allows also a deep look into the real contexts where each of the nominalisations occur and also the various contexts of the same nominalization, in this case ‘growth’ to see how and what collocates with it. Interestingly, results from the tool inform certain patterns whereby nominalisations deriving from verbs are seen to collocate with other nouns or prepositional phrases premodifying or post-qualifying the nominalized noun in the whole string with nominalisations. The complex semantic wrappings indicative of these nominalized segments like *wireless data **growth***, *the rate of wage **growth***, *the **growth** dreams of some developers*, *in better **growth** numbers*, *the unprecedented viral **growth***, *a huge area of **growth***. Figure 11 shows the context of *growth* and its concordance lines in the register of magazine in COCA.



Figure 11. Concordance lines of nominalisation 'growth' in COCA



The nominalised constructions deriving from an adjective or de-adjectival nominalisations can also be extracted, such suffixes as *-ity*, *-ness* for example, denoting *the status or the feature of being the adjective*. Concordances from COCA or BNC can throw further insight into the operationalisations of these nominalised segments which are found to collocate with other nouns or nominal groups or prepositional phrase with *of* post-qualifying the nominal and predeterminer *the*:

While Shields believes in the authenticity and continuity of the craftsman and does not interrogate craft masculinity as a social construct

Figure 12 illustrates the concordance lines of de-adjectival nominalised segments with *-ity* in COCA while Figure 13 similarly prescribes the window to the context of concordance line 16 when a click is made on its context.

Corpus of Contemporary American English																
SEARCH			FREQUENCY			CONTEXT			ACCOUNT							
(SHUFFLE)																
CLICK FOR MORE CONTEXT																
<div><input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></div>																
1	2016	ACAD	Human Rights Quarterly	A	B	C	1,72 Hence, debate in the US about the legal- ity of drone killing has been largely conducted in terms of the respective constitutional powers of								SHOW DUPLICATES	
2	2016	ACAD	Technology and Culture	A	B	C	, and his culture. " 50 While Shields believes in the authenticity and continu ity of the craftsman and does not interrogate craft masculinity as a social construct,									
3	2015	RIC	KerryonRev	A	B	C	Our mother was an only child, her parents long passed. The major ity of the attendees were former students and colleagues. The college sent over flowers,									
4	2015	RIC	FantasySciFi	A	B	C	N O O N found the discriminate r once in one in th e with ity of U to h's w arnhouse. T his tim e he had									
5	2015	RIC	LiteraryRev	A	B	C	to the Municipal Building to find out about his application for assistance and the possibl ity of the messenger job. Unable to pay for parking, he leaves his car									
6	2015	RIC	NorthAmRev	A	B	C	sting of the cold air filling their lungs. Conrad became aware of the proxim ity of his body to Daniel's -- the touching of thighs, the hands									
7	2014	RIC	SouthwestRev	A	B	C	laughs. " Now, Old friends. And parasailing is easy. The ambigu ity of parasailing gets me. " He sets the broom down, leaning his chin									
8	2013	RIC	NorthAmRev	A	B	C	not in Harlem'd met up with Dina when he'd gotten back to New ity of Dina, in the million miles of black asphalt highway Smyrna in time for									
9	2011	RIC	KerryonRev	A	B	C	were struggling beyond: the loss, and on top of that the impossibl ity of carrying out his last foolish, really foolish, wish. They regretted that									
10	2011	ACAD	CommCollegeJ	A	B	C	the opportunity structures of higher education. We begin with a discussion of the ambig- ity of postsecondary remediation, underscoring, in particular, the ov									
11	2010	ACAD	PalliativeMedicine	A	B	C	were most likely to experience psychological and symptom distress (increased pain, decreased qual- ity of life, difficulty with bowel functioning), heightened dep									
12	2010	ACAD	Lupus	A	B	C	dial patients were interviewed by telephone with a view to establishing content valid- ity of a Swedish version of the Quality of Life Scale (QoLS). In									
13	2010	ACAD	WarinHistory	A	B	C	a pioneer of aerial warfare, indicated the will of Chessa to entrust the responsibl- ity of its programme to one of the most important aviation experts. Having tal									
14	2010	ACAD	CulturalGeng	A	B	C	scaffolding of racial segregation in the US. David Delaney carefully demonstrates that the spatial- ity of racism was a central component of the social structure c									
15	2000	RIC	Analog	A	B	C	of the fact that every significant advance in biotechnology is seen by the vast major ity of horrified onlookers as a hideous perversion, we do make progress. We									
16	1999	MAG	HarpersMag	A	B	C	ethic kick in like a shot of oxytocin, the nurturance hormone. The plural ity of my customers are hard-working locals truck drivers, construction workers, even h									
17	1990	RIC	BlkMigrantsSouls	A	B	C	, accustomed to life among the less privileged, the same cars symbolized the ingenu ity of the poor. " They have not even built the new church. We									
18	1990	RIC	BkSFIslander	A	B	C	for you, it will involve travel and no little danger, and a possibl ity of great profit, just the thing, I believe, for one of your									

Figure 12. Concordance lines of nominalisations with *-ITY* in COCA

Corpus of Contemporary American English

SEARCH

FREQUENCY

CONTEXT

CONTEXT +

Source information:

Date	1999 (Jan)
Publication information	Vol. 298, Iss. 1784; pg. 37, 16 pgs.
Title	Nickel-and-dimed: On (not) getting by in America
Author	Barbara Ehrenreich
Source	MAG: Harpers Magazine

Expanded context:

about the table whose order I screwed up so that one of the boys didn't get his kiddie meal until the rest of the family had moved on to their Key Lime pies. That's the other powerful motivation I hadn't expected the customers, or " patients, " as I can't help thinking of them on account of the mysterious vulnerability that seems to have left them temporarily unable to feed themselves. After a few days at the Hearthside, I feel the service ethic kick in like a shot of oxytocin, the nurturance hormone. The plural *ity of* my customers are hard-working locals truck drivers, construction workers, even housekeepers from the attached hotel-and I want them to have the closest to a " fine dining " experience that the grubby circumstances will allow. No " you guys " for me; everyone over twelve is " sir " or " ma'am. " I ply them with iced tea and coffee refills; I return, mid-meal, to inquire how everything is; I doll up their salads with chopped raw mushrooms, summer squash slices, or

Figure 13. Context of *plurality* as a de-adjectival nominalisation in COCA

Nominalised constructions with **-ness** are searched in similar ways and results reveal showing the frequencies, contexts of situation and the wide variety of genres which are characteristic of such language patterns. Accordingly, the patterns with de-adjectival nominals with **-NESS** are strongly indicative of the *-of*-phrase being in the post-qualifying function. As Figure 14 illustrates, the nouns following *of* in the nominalised constructions can be coded as the agent that possess the features or quality of the adjective from which the nominalization is formed. For example, *the sheer richness of the painting* (concordance 22) must come from its congruent segment of ‘the painting is completely rich’ or *the slowness of her progress* (Concordance 41) being ‘her progress is slow’.

CONCORDANCE	YEAR	SOURCE	SEARCH	FREQUENCY	CONTEXT
21	2009	MAG	Ma	A B C	patients to learn more about their health status and options." according to Debra <b>ness</b> of the National Partnership for Women and Families. So what's left?
22	2008	MAG	AmerArtist	A B C	with a provocation, nothing, however, can detract from the sheer rich- <b>ness</b> of the painting, with its luxuriant fabrics, its wealth of foliage, and
23	2008	ACAD	VisualImpair	A B C	both the core curriculum and the expanded core curriculum. Comparative research on the effective <b>ness</b> of instruction in the expanded core curriculum pri-
24	2007	MAG	MotherJones	A B C	to sell \$325 million in shares to the Indonesian government, the provincial court acquitted <b>ness</b> of all charges. "I was amazed at how thoroughly the judges
25	2007	ACAD	IntAffairs	A B C	as follows: # The global resurgence of religion is the growing saliency and persuasive <b>ness</b> of religion, i.e. the increasing importance of religious beliefs, gras-
26	2006	RIC	FantasySciFi	A B C	the poor old phunt alone. "With a curious tenderness that belayed the gruff <b>ness</b> of his tone, he took off the old man's shoes and stowed them
27	2003	MAG	Money	A B C	we are, how little information we have to go on," says Debra <b>ness</b> of the Disclosure Project, a coalition of employees and consumers that is pushing health-
28	2000	RIC	Analog	A B C	strongly, like smoke and food and life. Every morning, the relentless gray <b>ness</b> of Ende pushed Linna closer to the brink of despair. The afternoons were late-
29	1997	NEWS	SanFrancisco	A B C	directing such serious films as "The Right Stuff" and "The Unbearable Light <b>ness</b> of Being," argues that money-making concerns alone aren't driving the d-
30	1996	MAG	RollingStone	A B C	in record and concert-ticket sales. That's a very expensive habit. # MIKE <b>NESS</b> of Social Distortion: # "I started getting interested in foreign films this year
31	1996	MAG	RollingStone	A B C	arranges her eyebrows into a sincere position. # PHOTO (COLOR): Mike <b>NESS</b> of Social Distortion # PHOTO (COLOR): Brian Green would just as soon
32	1994	MAG	Ebony	A B C	obligations: not to get lost in the sense of their specialness or the unique <b>ness</b> of the Harvard degree, but to always remember who they are. " "
33	1992	RIC	BkSF:GunsSouth	A B C	as if by trumpeting his middle name he could make up for the utter plain <b>ness</b> of those that flanked it -- finished writing his sentence before he looked up.
34	1992	RIC	BkSF:Sageary	A B C	saw that remaining sword flying out wide, and Drizzt, never doubting the effective <b>ness</b> of his strike, was moving forward before Twinkle ever connected. He
35	1992	RIC	BkSF:BreathSuspension	A B C	... institute artifacts. So the drug-dealer molded his product in the like <b>ness</b> of pottery shards. And the archeologist on-site made sure they got packed like a
36	1992	RIC	Bk:Songand	A B C	a rare white man, one with a good heart, one who understands the sacred <b>ness</b> of the animals. You will forever be our brother. " Tom felt a
37	1992	RIC	Bk:InterviewVampire	A B C	she told her charges cautionary tales. Her stories, however, of the good <b>ness</b> of good behavior collapsed before the threat of the sin they depicted. The truth
38	1991	SPOK	Chik_King	A B C	Welcome to Larry King Live. Tonight: He's been called the Eldest <b>ness</b> of the Food and Drug Administration, FDA Commissioner David Kessler creates color
39	1991	RIC	BkSF:DreadPhemara	A B C	after all. He started forward, then stopped. His eyes swept the dark <b>ness</b> of the old house, wondering if he might be walking into some sort of
40	1991	RIC	BkSF:SadersKeep	A B C	" "Cold look?" Liriel repeated softly, puzzling at the strange <b>ness</b> of Jobber's words. "Leave it," Jobber growled. "Just
41	1991	RIC	Bk:TrueColors	A B C	defended herself, sick over Cy's reminding of her physical needs and the slow <b>ness</b> of her progress. "Blake, don't push. You know I'd
42	1991	ACAD	CommCollege	A B C	for more federal legislation beneficial to community colleges, the need to improve the cohesive <b>ness</b> of public community colleges within each state, the co-
43	1990	RIC	Bk:PlainsPassage	A B C	far to the south the only place that came close to matching the profound rich <b>ness</b> of the Ice Age steppe the great abundance and variety of animals shared

Figure 14. Context of **plurality** as a de-adjectival nominalisation in COCA

Further concordance lines from COCA can also be taken from various years of media publications with real contexts being given for nominalisations with **-ness** (See Figure 15).

he didn't see what harm it could do to treat Clair with a little extra fond**ness** and leniency - *Virginia Quarterly Review*

And when I walked through the tinted-glass doors of the cool, cave-like computer lab into the warmth and bright**ness** of the school's courtyard, how satisfying to hear that familiar voice from the oaks or the wisteria trained on a trellis overhead. \**Southern Review*, Vol 15, Issue 1 2015

It was less sad than Jill had feared, the empt**iness** of the place more peaceful and powerful. *Ploughshares*, 2015, 41 (2)

With that Clair found him there, standing at the edge of the tank. Her face initially registered surprise but soon deepened into kind**ness**. "I didn't think you were going to make it - *Virginia Quarterly Review* 2015, 91-3.

11	2016	ACAD	Journal of Insurance Issues	A B C	tionally, the life business generally is longer-term in nature than the property-casualty busi- <b>ness</b> . Further, we ultimately document empirically that whik-
12	2016	ACAD	Early American Literature	A B C	; the Parisians are unwilling to rise with the sun, and their "lazi- <b>ness</b> " impacts perceptions of their economy's stability. Franklin attunes his economic ar-
110	2010	ACAD	TheologStud	A B C	# (n51) Of course Newman's contribution to the emergence of historical conscious <b>ness</b> within Catholicism can not be underestimated. # (n52) Joseph Ri-
111	2010	ACAD	Trauma	A B C	Y-CHIRRP a single children's hospital in Glasgow presenting problems in terms of the representative- <b>ness</b> of the data. Finally, the Health and Safety Exec
112	2010	ACAD	CulturalGeog	A B C	and intensified anxieties regarding the deleterious effects of alcohol on Indians, and how drunken- <b>ness</b> might trouble their moral upliftment and event.

Figure 15. Context of de-adjectival nominalisations with **-ness** in COCA

The COMPARE function is also used to compare nominalisations as near-synonyms. Two synonymous nominals like *development* and *growth* are put in the COMPARE box, and cases showing collocates with either



or both of these words are shown. For instance, collocates of *growth* with *love* is found rare, and still rarer with *development*. In fact, no such case throws up in the search.

The figure consists of two screenshots of the British National Corpus (BNC) interface. The top screenshot shows the search results for the word 'development'. The bottom screenshot shows the search results for the word 'growth'.

British National Corpus (BNC)			
SEARCH	FREQUENCY	CONTEXT	CONTEXT +
(SHUFFLE)			
CLICK FOR MORE CONTEXT			
1. BNC	W_at_posit_jaw_edu	A B C	...app maintenance and then a Profit share analysis of the development of the sales of profit in terms of two statistical components. One is the
2. G3	W_at_humanities_arts	A B C	...payments. This is a procedure founded upon a general estimate of the rate of development of distinct features. In other regions, e.g. Dorset or North Devon, a
3. BNC	W_at_posit_humanities_arts	A B C	...in the period 1924-4, can be summarised as being directed towards the development of a common rate of investment and growth as opposed to maximum
4. G3	W_at_posit_soc_sci	A B C	...that occupy temporary ponds, such as the European waterbird pool. The rate of development of aquatic habitats depends upon several factors, including water
5. G3	W_at_posit_soc_sci	A B C	...much disputed, argues comments, possible variables such as the rate of development of community services, the aging of the population, economic uncertainty, a
6. H1	W_posit	A B C	...Such warnings offer the potential for sizeable capital gains if property values rise. The development of new roads and carrying capacity has offered a further impetus
7. G3	W_at_posit_jaw	A B C	...automations for Body, in 3.1 The Science Case 4 The Science Case 4 The Science Case 4 The Science Case 4 The Science Case 4 The Science Case 4 The Science Case 4

  

British National Corpus (BNC)			
SEARCH	FREQUENCY	CONTEXT	CONTEXT +
(SHUFFLE)			
CLICK FOR MORE CONTEXT			
1. G3	W_at_soc_sci	A B C	...the World Health Organization by the World Health Organization and the Growth of Love, Harold Wertheim's theory in which the material deprivation is postulated
2. ACT	W_at_posit_humanities_arts	A B C	...On the growth of love, the author of a book, Child Care and the Growth of Love, which popularised the theory of Maternal Deprivation, the
3. G3	W_at_posit_humanities_arts	A B C	...The constant flow of the story depends in part on the contrasting yet inevitable growth of love between the young man and girl, and of his physical strength and
4. ABC	W_posit	A B C	...In cases of serious children, the collapse of love, and the growth of responsibility and indifference to children, will ultimately lead to the end of the
5. ABC	W_posit	A B C	...courtship and the first years of their marriage, others still wonder what is the growth of love, more, in consequence, are a continuation of the same

Figure 16. Some concordances with 'development' and with 'growth' in BNC

Corpora like COCA and BNC also allow comparison made into the patterns of language used in different registers. Occurrences and percentages corresponding to each of the genres where the search word or phrase are examined can be shown in nicely-presented tables, providing interesting information about the features of any language patterns (Figures 17-18).

The screenshot shows the British National Corpus (BNC) interface with the search results for the word 'behaviour'. The results are displayed in a table with columns for ID, Corpus, Register, and Context.

British National Corpus (BNC)			
SEARCH	FREQUENCY	CONTEXT	CONTEXT +
(SHUFFLE)			
CLICK FOR MORE CONTEXT			
1. ANS	W_at_posit_jaw_edu	A B C	importance of ascertaining pupils' views both generally and more specifically in relation to the development of a behaviour policy by the school, and to instil a sense
2. G3	W_at_nat_sci	A B C	Behaviour patterns can not be developmentally divided into inherited and learned behaviour, for the development of every behaviour pattern is influenced by both
3. G3	W_at_nat_sci	A B C	of experience has been studied in many systems. It is undoubtedly important in the development of behaviour, and makes sense from the point of view of fitting in
4. G3	W_at_nat_sci	A B C	behaviour develops by an interaction of genes and environment. An account of the development of behaviour within an individual would have to mention a series
5. G3	W_at_nat_sci	A B C	of experience instinctive. Of course he did not deny that experience does influence the development of some behaviour patterns, and he accordingly divided beha
6. G3	W_at_nat_sci	A B C	ladybirds. It is possible to discover what factors are not needed for the normal development of a behaviour pattern by isolating young animals from hypothetically
7. G3	W_at_nat_sci	A B C	Behaviour patterns can not be developmentally divided into inherited and learned behaviour, for the development of every behaviour pattern is influenced by both
8. G3	W_at_nat_sci	A B C	(1986) has recently started to look at the contribution siblings make to the development of aggressive behaviour. Naturalistic studies in the homes of families with a
9. EDH	W_at_soc_sci	A B C	necessity of human contact and care for the development of social behaviour and for the development of fundamental human faculties. Consider how Sheldrake fir
10. EDH	W_at_soc_sci	A B C	not instinctive but learnt, indicating the necessity of human contact and care for the development of social behaviour and for the development of fundamental hum
11. J14	W_at_soc_sci	A B C	possesses. # (R) Sex # Boys are constitutionally more predisposed to the development of aggressive behaviour than are girls (Wolff, 1983a). There may also
12. G3	W_at_posit_humanities_arts	A B C	3The genetics and development of behaviour # 3.1 The principles of genetics # inheritance has been the subject of
13. HHY	W_at_posit_soc_sci	A B C	3 # the emergence of agency and self efficacy and individual differences in the development of self-regulated behaviour and attitudes. The contextualising of these





Figures 17-18. Concordances, occurrences and percentages of 'development' in BNC

## V. Conclusion

With the main goal of teacher as researchers being established which is to search innovative ways to improve language teaching and learning, knowledge on language patterns driven by large data size is thus important to really inform language teaching, thereby enhancing language learning. Without any shadow of a doubt, all teaching transformations require that new techniques and/or specific pedagogic technologies be incorporated in computer-assisted ways and by means of web-based resources and tools with we are to keep up with the pace of such development.

With more and more teachers as researchers probing into the depth of language in use with the aid of corpus-based web tools, the future is not far away when increasing authenticity and practicality is imported into the classroom. Hopefully, minor case studies as presented in this paper are of some relevance to the on-going endeavor for bridging the gap between corpus linguistics as a sub-field of linguistics and corpus-based studies driven by web resources as toolkits for further explorations into language patterns for enhanced language pedagogy.

## REFERENCES

- [1]. Ahmadi, M. R. (2018). The Use of Technology in English Language Learning: A Literature Review. *International Journal of Research in English Education*, 3 (2), 115-125.
- [2]. Baker, P. (2010). Corpus methods in linguistics. In L. Litosseliti (ed.), *Research Methods in Linguistics* (pp. 93–113). London: Continuum.
- [3]. Braun, S., Korn, K. and Mukherjee, J. (2006) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*. English Corpus Linguistics, Vol 3.
- [4]. Diakou, M (2015) Using information and communication technologies to motivate young learners to practice English as a foreign language in Cyprus. *Jaltcalljournal*, Vol. 11, No.3, pp. 283–292.
- [5]. Erenchinova, E. and Proudchenk, E. (2017) Advantages of WEB resources use in learning process. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, Volume 97, 103-108.
- [6]. Geladze, D. (2015) Using the Internet and Computer Technologies in Learning/Teaching Process. *Journal of Education and Practice*, Vol, 6, No.2, 67-70.
- [7]. Kamba, M. A. (2007) The Internet as a Tool for Interactive Learning, Teaching and Research: Nigerian Experience. IMCL Conference.
- [8]. McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Methods, Theory and Practice*. Cambridge: Cambridge University Press.
- [9]. Ngula, R. S. (2018). Corpus Methods in Language Studies: Perspective on Conducting and Reporting Research in the Humanities, 205-223.
- [10]. Pim, C. (2013). Emerging technologies, emerging minds: Digital innovations within the primary sector. In G. Motteram (Ed.), *Innovations in Learning Technologies for English Language Teaching* (pp. 15–42). London: British Council.
- [11]. Son, J.-B. (2011). Using Web-Based Language Learning Activities in the ESL Classroom. *International Journal of Pedagogies and Learning*, 4(4), 34-43.
- [12]. The British National Corpus (BNC)
- [13]. Contemporary Corpus of American English (COCA)
- [14]. The American Facebook Page.