

How to Improve Your Academic Writing Using Language Corpora

Author

Enago Academy

Post Url

<https://www.enago.com/academy/how-to-improve-your-academic-writing-by-using-corpora/>



No matter how brilliant a researcher you are, you must be able to write about your research effectively to have any impact on the scientific world. Unfortunately for most of us, research and writing are two very different skills. Even the most talented researchers may struggle when it comes to writing clearly and concisely about their work. The burden is doubled for non-native English speakers. While English is widely accepted as the global language of science, it is also a tricky and difficult language to learn. What is the difference between “put on” and “put off”? Do you “take” a sample or “make” a sample? Where can you go when you need help with English writing? One little known, underused source of help for academic writing is language corpora. In this article, we

will talk about how you can take advantage of this resource to improve your writing and increase your confidence with English.

enago
Author First. Quality First

STARTING \$12

Enago Plagiarism Check with AI-based Grammar Correction

World's leading plagiarism check with powerful AI grammar check for your writing, delivered in just about 5 minutes.

91 Billion+ Web pages
102 Million+ Pages crawled daily
82 Million+ Scholarly articles
37 Hundred+ Open access repositories

CHECK NOW

Powered by Turnitin

What is a Language Corpus?

A language corpus is a collection of electronic text used for research purposes. Language corpora were originally created by researchers, usually linguists, for research purposes. Some popular corpora include the [Corpus of Contemporary American English \(COCA\)](#), [Corpus of Historical American English \(COHA\)](#), [Google Books Ngrams viewer](#), [Michigan Corpus of Academic Spoken English](#), [Hyper Collocation](#), and more. These corpora offer a searchable collection of English used by native speakers in different contexts. In English language classes, they are often [used as a tool by teachers](#) who want to show their students how a word is used in real life by native speakers.

What is the difference between a corpus and a dictionary? Why would a non-native English speaker turn to a corpus instead of a dictionary for answers? First of all, while a dictionary can define a word for you, it often does not include many usage examples. The word “extract” means “to remove or take out.” But if I need to know how to explain a physical action I took in my research, will I say “extract to” or “extract from”? A dictionary probably cannot answer that question, but a language corpus can.

Familiarizing yourself with some simple corpus search functions will make a new range of tools available to you. Many corpora allow searches for synonyms and different word forms. For example, you could search for the verb form of “extract” using COCA and return “extracts,” “extracting,” “extracted,” and “extract.” You could also select “collates” for your search string and return a list of words that are frequently found together with the word “extract.” Clicking the “help” icon will offer you a variety of search function methods. For example, if you type in [=extract] you can find a list of synonyms for the word such as remove, separate, get, fetch, and so on.



Another advantage of language corpora is that they are updated more frequently than dictionaries. A search in Webster’s dictionary in early 2019 would not have returned a

result for the term “bioabsorbable.” But the word has been in use and popularized thanks to new advances in technology that were [presented in 2019](#). The word [was officially added](#) to Merriam Webster’s in the middle of 2019. If you were looking for examples of how to write using this word, corpora would be there to provide you with examples of contemporary use.

How Do I Use Language Corpora?

Learning to search on different language corpus tools can seem confusing at first. But don’t worry- it gets easier quickly. Now let’s look at how to choose a corpus and how to search for different words on these sites to get useful results.

You should choose your language corpus depending on what your goal is. If you are looking for how to use a word that is not specific to your discipline, then COCA will be a great place to start. Let’s say you want to know if you should say “extract to” or “extract from.” You can click on the link to COCA above and enter the term “extract to” in the search bar. Then you will click “find matching strings.”


Corpus of Contemporary American English        

SEARCH FREQUENCY CONTEXT OVERVIEW


[List](#) [Chart](#) [Collocates](#) [Compare](#) [KWIC](#)

extract to [POS]

Sections [Texts/Virtual](#) [Sort/Limit](#) [Options](#)

 (HIDE HELP) NOT LOGGED IN

Same corpus - new location: English-Corpora.org

 Download the corpus for offline use

The Corpus of Contemporary American English (COCA) is the only large, genre-balanced corpus of American English. COCA is probably the most widely-used corpus of English, and it is related to many other corpora of English that we have created, which offer unparalleled insight into variation in English.

The corpus contains more than 560 million words of text (20 million words each year 1990-2017) and it is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts.

Click on any of the links in the search form to the left for context-sensitive help, and to see the range of queries that the corpus offers. You might pay special attention to the [comparisons between genres and years](#) and the (new) [virtual corpora](#), which allow you to create personalized collections of texts related to a particular area of interest.

[More help files](#)

When we perform this search for “extract to,” we return only 52 uses, while “extract from” returns 233.



Corpus of Contemporary American English        

SEARCH FREQUENCY CONTEXT HELP

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

[iWeb](#) EXTRACT TO

| | <input type="checkbox"/> CONTEXT ALL FORMS (SAMPLE): 100 200 500 | FREQ | |
|---|--|------|---|
| 1 | <input type="checkbox"/> EXTRACT TO | 52 | <div style="width: 100%; height: 10px; background-color: #007bff;"></div> |
| | | | 0.221 seconds |

Corpus of Contemporary American English        

SEARCH FREQUENCY CONTEXT HELP

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

[iWeb](#) EXTRACT

| | <input type="checkbox"/> CONTEXT ALL FORMS (SAMPLE): 100 200 500 | FREQ | |
|---|--|------|---|
| 1 | <input type="checkbox"/> EXTRACT FROM | 233 | <div style="width: 100%; height: 10px; background-color: #007bff;"></div> |
| | | | 0.268 seconds |

We can click on “context” to see exactly how it is used. Based on this search, we will decide that “extract from” is the correct word form to use.

| Corpus of Contemporary American English | | | | | | | | | | | | | |
|---|------|-----------|-----------------|---------|---|----------|---|--|--|--|--------------------------|-----|-----------------|
| SEARCH | | FREQUENCY | | CONTEXT | | OVERVIEW | | | | | | | |
| (SHUFFLE) | | | | | | | | | | | | | |
| CLICK FOR MORE CONTEXT | | | | | | | | | | | <input type="checkbox"/> | [?] | SHOW DUPLICATES |
| 1 | 2017 | SPOK | NBC: Today Show | A | B | C | . They're incredibly soothing. They help to depuff. They've got licorice extract to brighten and they've got things that will help smooth out the skin but | | | | | | |
| 2 | 2015 | MAG | NaturalHist | A | B | C | fractions of different-size molecules). He then timed how long it took for each extract to take effect. When mice were paralyzed, they would let go of the | | | | | | |
| 3 | 2014 | MAG | Esquire | A | B | C | to slow the damage caused by the sun and toxins. This cream uses ginseng extract to stimulate the skin, so you see benefits daily and not just over time | | | | | | |
| 4 | 2014 | MAG | Prevention | A | B | C | , that's not an issue. An alternative remedy: Apply grapefruit seed extract to the nail. Eating a daily clove or two of garlic, which has | | | | | | |
| 5 | 2014 | NEWS | Atlanta | A | B | C | stick butter, softened 1 cup dark brown sugar 1 egg 1 teaspoon pure vanilla extract To make the caramel apples: Generously grease an extra-large casserole dis | | | | | | |
| 6 | 2014 | NEWS | Atlanta | A | B | C | stick) unsalted butter, melted, plus more for the pan 1/2 teaspoon lemon extract To make the topping: In a small bowl, stir together the all-purpose flour | | | | | | |
| 7 | 2013 | FIC | TexasRev | A | B | C | fill water tanks on Overlook Ridge when electric service became dicey, used creosote bush extract to banish the plague of flies, worked frantically to divert Slugg | | | | | | |
| 8 | 2013 | MAG | Essence | A | B | C | permanent hair color system uses natural and certified organic ingredients like shea butter and comfrey extract to provide moisture and promote shine, and vit | | | | | | |
| 9 | 2013 | NEWS | OrangeCR | A | B | C | more than 3,000 lab tests. These textured wipes are made with Vitamin C-rich grapefruit extract to cleanse, exfoliate and brighten skin. They tingle a bit when yc | | | | | | |
| 10 | 2013 | ACAD | JAnimalPlantSci | A | B | C | total phenolics: The total phenolic content was determined by adding 0.5 ml of the extract to 2.5 ml, 10% Folin-Ciocalteu's reagent (v/v) and 2 ml sodium | | | | | | |
| 11 | 2012 | MAG | Prevention | A | B | C | says. # TRY # The trials used 250 mg tablets of Kyolic aged garlic extract to standardize the dose. " But it's always better to eat the real | | | | | | |
| 12 | 2012 | MAG | ParentingSchool | A | B | C | , until set but still pale in color. # Maple Add 2 tsp maple extract to the dough. Sprinkle cookies with raw (turbinado) sugar before baking. | | | | | | |
| 13 | 2012 | MAG | Prevention | A | B | C | on hair. Twisted Sista Leave-in Conditioner Spray (\$5; drugstores) has lychee extract to prevent UV damage; caffeine (from coffee) stimulates follicles in Barista E | | | | | | |
| 14 | 2010 | MAG | Cosmopolitan | A | B | C | (or a drop of Pama Liqueur) to one, a drop of mint extract to another, and a splash of coffee to another -- they'll all taste | | | | | | |
| 15 | 2009 | MAG | SouthernLiv | A | B | C | as directed. Makes about 2 cups. Vanilla Latte Buttercream Frosting: Increase vanilla extract to 2 tsp. Proceed with recipe as directed, beating in 1 Tbsp. instant | | | | | | |
| 16 | 2009 | MAG | Essence | A | B | C | the butter into the dry ingredients mixture until it clumps when pinched. Add vanilla extract to the cream and slowly add the two to the dry ingredients while tu | | | | | | |
| 17 | 2009 | MAG | Prevention | A | B | C | may need to call in the special forces- ingredients such as vitamin K and licorice extract to minimize darkness or caffeine to reduce swelling, which are available | | | | | | |
| 18 | 2008 | MAG | Shape | A | B | C | (\$460; all at saks.com), which is infused with white willow leaf extract to improve the elasticity of slack skin. YOU WANT TO RESTORE MOISTURE Try The | | | | | | |
| 19 | 2007 | SPOK | CBS_SatEarly | A | B | C | And I just add that and some sugar. And a little bit of vanilla extract to just bring out the flavor, and then you put this over the heat | | | | | | |
| 20 | 2007 | MAG | Prevention | A | B | C | several animal studies. When researchers at the University of Illinois at Urbana-Champaign fed rosemary extract to rats exposed to dimethylbenzanthracene, a | | | | | | |
| 21 | 2007 | MAG | Prevention | A | B | C | least 0.3% hypericin (the active phytochemical) per capsule or 300 mg of the extract to be taken three times daily. Warning: St. John's wort has been | | | | | | |
| 22 | 2007 | MAG | VegTimes | A | B | C | options * Earth Science Clarifying Herbal Astringent uses zinc to tame oiliness and witch hazel extract to naturally tone and remove those last traces of makeup. | | | | | | |
| 23 | 2007 | MAG | SatEvenPost | A | B | C | the University of Toronto. 1908 German scientist Georg Zuelzer develops the first injectable pancreatic extract to suppress glycosuria; however, there are extrer | | | | | | |
| 24 | 2007 | MAG | Shape | A | B | C | help exfoliate dead skin and clear pores, it has antioxidant vitamin E and grape extract to fend off signs of premature aging. Downie calls it " gentle yet luxurious: | | | | | | |
| 25 | 2007 | MAG | Shape | A | B | C |) Los Angeles dermatologist Howard Murad, M.D., has turned to a super-concentrated pomegranate extract to boost the efficacy of this sunscreen. In an indepe | | | | | | |
| 26 | 2007 | MAG | Shape | A | B | C | you want to " tan. " Besides delivering color, this lotion has cocoa extract to intensify your hue, aloe vera for softness, and antioxidant vitamin E to | | | | | | |
| 27 | 2007 | NEWS | USAToday | A | B | C | # Just because drinking tea might be good for you doesn't mean adding tea extract to cereal and other foods or to dietary supplements is beneficial. " A lot | | | | | | |
| 28 | 2006 | NEWS | WashPost | A | B | C | muffins. # In a food processor, add the butter, honey and vanilla extract to taste and process until smooth. Serve with the warm muffins. # *NOTE | | | | | | |
| 29 | 2005 | MAG | Shape | A | B | C | Defense Foaming Bath (\$6) contains vitamins A and C, ginseng and white-tea extract to nourish skin and fight skin-aging free radicals while you're in the tub. | | | | | | |

For more discipline-specific words, you can try the Michigan Corpus of Academic Spoken English ([MICASE corpus](#)), which offers some limited examples. The advantage of Michigan's tool is that [you can search by discipline or type of academic event](#). If you are writing to prepare for a specific type of event or branching into a new part of your field, this tool can be particularly helpful for you.

You may also be wondering about the differences between American and British English. Don't worry- there are corpora to help you with those searches too. The [BYU Corpus site](#) has links to British English and American English corpora, and you can search and compare to see what terms or phrases are used in one style over the other. Should we say "in hospital" or "in the hospital"? A search of the corpora shows that Americans favor "in the hospital," while British English speakers simply say "in hospital."

A Few Notes of Caution

You may be very excited to begin using this new tool. You should be! Language corpora can be extraordinarily helpful in providing you with real-world examples of language that you would have difficulty finding otherwise. Dictionaries and Google searches do not provide nearly the amount of detail and context that corpora do. However, there are still some [points of caution to keep in mind](#) when relying on corpora to improve your writing. First, corpora don't tell you what is correct and incorrect. They simply tell you what usage is common. You can use corpora to improve your writing, but you may need to dig deeper and compare your data from corpora with other sources.

That said, language is a funny thing. What is key to remember is that language is about communication. When you seek out how to use certain words, real-world examples are a great tool that can give you a new and deeper level of understanding of the words themselves. For that reason, language corpora are a great tool to have in your toolbox when it comes to improving your academic writing.

Do you use language corpora to help you in academic writing? Which corpus do you find most helpful? What are some other good resources for ESL writers to improve their academic writing? Let us know in the comments below!

Cite this article

Enago Academy, How to Improve Your Academic Writing Using Language Corpora. Enago Academy. 2019/12/27. <https://www.enago.com/academy/how-to-improve-your-academic-writing-by-using-corpora/>