

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.



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 Historical 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 <u>presented in 2019</u>. The word <u>was officially added</u> 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?

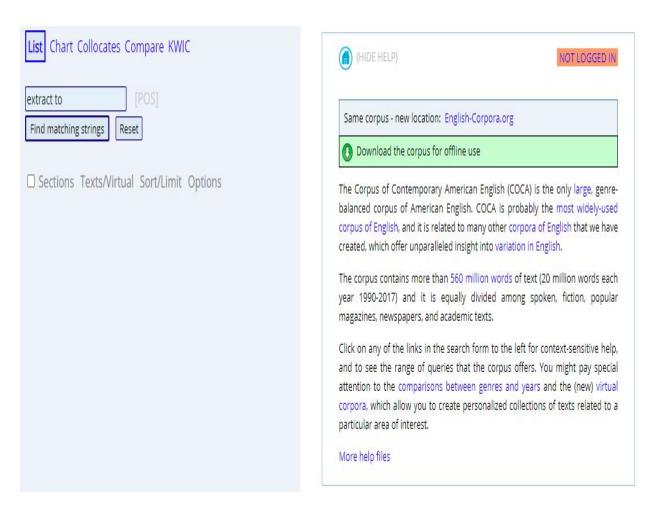
Learning to search on different language corpus tools can seem confusing at first. But don't worry- it gets easier quickly. Now let's look at how to choose a corpus and how to search for different words on these sites to get useful results.

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."









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

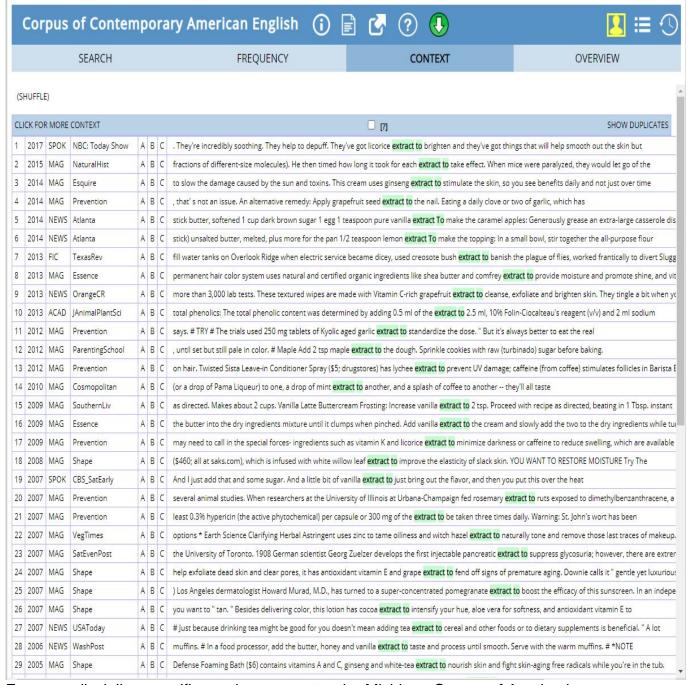






We can click on "context" to see exactly how it is used. Based on this search, we will decide that "extract from" is the correct word form to use.





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/

