

Corpus of Contemporary American English (COCA)

<http://corpus.byu.edu/coca/> Login: ohiostatewritingcenter3@gmail.com Password: wr1+1ng

English is a difficult language to master. Many rules of English grammar break down on application, and the common uses of words are often learned through direct observation. Therefore many writers have trouble telling whether something *sounds right* to Americans.

The Corpus of Contemporary English (**COCA**) collects 520 million English words as they have been used in speech, in writing, on TV, in academic writing, and other areas between 1990 and 2015. COCA has many potential uses, including: (1) researching common American idioms (unique expressions); (2) double-checking prepositions and verbs; (3) comparing styles between spoken and academic American English; (4) finding the right word for each medium

There are more uses that you can discover with independent research. All of these uses begin with a basic search, accessible at <<http://corpus.byu.edu/coca/>>. The search bar will appear will appear in the upper-right of the screen.

1. Researching common American idioms

Common American idioms can be researched with the default search for COCA, called a “List” search.

- a. **Type a word** or phrase into the search bar, and press Enter.
COCA will show a list of matches in blue all-caps hyperlinks.
On the right, the results will show the **frequency** of that word or phrase in COCA.
☆ *For example, “helpmeet”*
- b. **Select the word** or phrase hyperlink that best fits your interests.
The selection will show every usage of that word or phrase within COCA.
By default, COCA will display (from left to right) the year, the medium, the publisher, and a selection from that **source** that shows the word or phrase.
☆ *For example, “2013 MAG AmSpect”*
- c. **Click** on the year, medium, or publisher to see more from that selection.
COCA will display the publication information about that source, as well as an **extended quote** containing the original selection.

COCA can search for American idioms that are only known in part. For example, if you are searching for an idiom about the American flag that begins with “star” and ends with “banner,” you can search COCA for “star * banner”. The asterisk [*] sign is a “wildcard” for search, so that any text that fits between “star” and “banner” will appear in COCA’s results. In this example, the most common result is “Star-Spangled Banner,” with 276 hits in COCA, compared with 1 hit for “Star-and-Cross-Spangled Banner.”

2. Double-checking prepositions and verbs

Common combinations of prepositions and verbs can be researched with one additional step on the default search.

- a. **Type a word** into the search bar, SPACE, and select [POS] next to the search bar. COCA will replace [POS] with a drop-down menu listing **parts of speech**.

☆ For example, “learn_”

- b. **Select the part of speech** that would fit your search term.

If you typed in a verb and need to know a preposition, then select “prep.ALL”

☆ For example, “learn [i*]”

- c. **Press Enter.**

COCA will show a list of matches in blue all-caps hyperlinks.

On the right, the results will show the **frequency** of that word or phrase in COCA.

☆ For example, “learn about”, 5436; “learn from”, 5204; “learn as”, 60

The most frequent results are most common among English-speaking Americans.

3. Comparing styles between spoken and academic American English

COCA’s frequencies can also show the relative appropriateness of a word for each section, such as the unpopularity of “stuff” in academic English. COCA’s frequencies can also show changes in popularity, such as the rising popularity of “stuff” between the earliest section and the latest.

- a. **Type a word** in the search bar, and then select “Sections” and check the box next to “Sections.” Press Enter.

☆ For example, “stuff”

COCA will show the frequencies of the word according to each section

☆ For example,

Spoken, 22584; Academic, 9098; 1990-94, 9230; 2010-2015, 15122

- b. **Click the number** of any section.

COCA will show the “List View” of that word within that section.

These results can’t show what is the right appropriate without some context, but COCA can show which terms appear more frequently in spoken or academic American English.

4. Finding the right word for each medium

COCA can help identify what patterns of use are appropriate in different media, including academic writing. For example, COCA can show that “stuff” is more commonly used in Academic Geology and Social Science publications than in Philosophy publications.

- a. **Type a word** in the search bar.

☆ For example, “stuff.”

- b. **Select a medium** listed under “Sections.” Press Enter.

☆ For example ACAD: Philosophy and ACAD:Geog/SocSci

COCA will show the “tokens” or number of the word according to each medium

☆ For example, “Philosophy” has 128 “tokens”; “Geog/SocSci,” 324

These results also can’t show which words are always appropriate, but they can indicate what is more or less common in any of the designated fields.