**COCA (Corpus of Contemporary American English) Basics**

1. **Explain COCA: COCA is online concordancer that uses Corpus of Contemporary American English**

***Corpus: an electronic database of written texts or transcriptions of speech (similar to text files on the internet. Difference is that it only has “words”, not images or videos) Simply put, it is a large collection of texts. COCA is a corpus of texts in contemporary American English***

***Concordancer: A tool to search through and analyze a corpus (similar to google search engine, or “ctrl + F” function of your computer. However, a concordancer is much more sophisticated than google. The most common usage of concordancing for ESL learners is to see if certain words (or strings of words) are actually used or how frequently they are used.***

1. **Getting started**

**Go to COCA at BYU—google it and you’ll find it.**

1. **Basic functions**

**A. Display Options**

* 1. **LIST: Show** a list of word(s) or combination of words (ranked according to their frequency)

**spring break**

* 1. **CHART: Show a chart comparing frequencies of a word in differentgenre or time.**

**cool vs. suffrage**

* 1. **KWIC: Show the key word(s), i.e. search word(s), in context.**
	2. **COMPARE: Compare two words according to their frequencies (just generally or with a certain collocate).**

**hot debate vs. heated debate**

**B. Types of queries (Search string)**

1. **Word: a search word or phrase**
2. **Collocates: a word (not a phrase) that occurs within up to 10 words before / after the search word(s) You can choose the collocation range by clicking two little boxes next to the COLLOCATE box.**
3. **POS LIST: List of “parts of speech”. This function is used when you don’t know the exact word (or a collocate of the word) you are looking for, but you know which part of speech (e.g., noun, verb, or adjective) you want. This is also called “wildcard” search. Place the cursor in the word or collocates box and click the drop-down arrow and select a part of speech that you are looking for.**
* **For example, if you don’t know which preposition to use in a sentence like “I am going to New York \_\_\_ Spring Break”, you may try…**

**[WORD]: Spring break**

**[COLLOCATE]: [i\*] (This will appear if you click the “prep ALL” option in the POS LIST)**

**From the search result, you can conclude that “on” is most frequently used with *spring break*.**

**COCA (Corpus of Contemporary American English) Exercises**

**A. Learning collocations with COCA**

**Step 1: Look at the example sentence below.** Answer **questions (a) through (c). Then circle the word that you should use in your paper.**

|  |
| --- |
| I am *fully / totally* aware of the problem.  |

1. In which genre is “totally” most frequently used?
2. In which genre is “fully” most frequently used?
3. Thus, which word would you use in your paper?

**HINT) You should use the chart display option.**

**Step 2: Using COCA, find the more frequently used collocate for the word “technology” in the sentence below.**

|  |
| --- |
| I’m studying the *utilization / application* of modern **technology** in class. |

**Answer: \_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**HINT) You should use the compare display option with collocate search string.**

**Step 3: Look at the following sentences. Which one is correct? Use the KWIC option in COCA to find the answer.**

 **1) I am looking forward to meeting you in class.**

 **2) I am looking forward to meet you in class.**

 **HINT) You should use the kwic display option.**

**Step 4: Look at the example sentences below. The underlined word in each sentence is an awkward collocate of the word in bold. Using COCA, find better collocates and revise the sentences.**

|  |
| --- |
| 1. **I hope to *succeed* the goal. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
2. There has been a *hot* **debate** over the issue. **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**3) He *firmly* recommended this place. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |

**HINT) You should use list display option with a wildcard (v\*, adj\*, adv\*) collocate or a synonym collocate search string.**