Using a Corpus to Help Students Learn by Themselves

A tool that can be very useful for instructors, tutors, and students is a corpus. In linguistics, a corpus refers to a collection of written and spoken texts that are compiled into a database. These written and spoken words are categorized into different genres (academic English, news, spoken English, etc.) allowing for anyone to search how language from speakers of English is used in real world situations and the frequency it is used.

While a dictionary is a good way to look up what a word means and to see the different definitions of a word, a corpus shows how the word is actually used. It shows how speakers and writers use the word in different contexts. This allows an English language learner to see copious amounts of examples for any given word in a detailed context and for them to see the pattern of usage for themselves. Just as a dictionary is a tool for understanding new words, a corpus presents words in detailed and real contexts. It allows the learners to see the "proper" or "appropriate" use of a word they may understand but feel that they may not be using correctly.

One popular corpus is the Corpus of Contemporary American English or COCA. To use it, sign up for a free account here: https://www.english-corpora.org/coca/. Once you sign up, you can get started searching any word or phrase within the corpus. After signing up, click on "WORD" and type in any word you would like to learn about. For example, you could type "homework". Once you do that, you will see the results. It should look like this:



Notice that you can see how frequently the word appears in different genres of English (blog, web, TV, spoken, etc.) This can help with figuring the appropriate use of the word in different situations, something English language learners have some difficulty with. Also, you can see the collocates (words that frequently appear together). This can help learners to chunk words together which improves reading fluency and diction in writing. Notice that this can be used for ALL levels of learners. For higher level students, they should get their own account and use it as a resource. For lower level students, instructors and tutors can guide the student to "discover" the use of a word by themselves which may result in increased motivation in learning

and in remembering how to use the word in the future. COCA has a plethora of functions – there's even a pronunciation/intonation function, courtesy of YouGlish. To get a quick introductory overview of how to use its most basic functions, click on this link: https://www.youtube.com/watch?v=yIEgBmoViAc.