

5. An awareness of and an ability to use Corpus analysis tools

An introduction to AntConc

Before we begin, discuss the following questions with a partner.

1. What is a corpus?
2. Have you ever used one?
3. Do you know of any which are used to better understand academic writing?
4. Have you heard of AntConc? Do you know what it does?
5. Can you think of any ways in which concordance software programmes can help you analyse specific genres?

Key information

The following information has been adapted from 'A guide to using AntConc'.

A **Concordance** is a list of target words extracted from a given text, or set of texts, often presented in such a way as to indicate **the context in which the word is used**. This format of presenting information is called 'KWIC': Key Word In Context. Concordance software can usually extract and present other types of information too, e.g. identifying the words that most commonly appear near a target word (its 'common collocates').

AntConc is a freeware concordance program developed by Prof. Laurence Anthony, Director of the Centre for English Language Education, Waseda University (Japan).

The AntConc programme is available from

<http://www.antlab.sci.waseda.ac.jp/software.html>

You can either download it onto your computer or run it from the website.

Task

You are now going to work with a number of texts to see how AntConc could help you in the classroom.

Step 1: Choose a number of texts (I would suggest 3-5). The texts can be from websites, articles or student essays. Save these texts as .txt (I will show you how to do this, if you are unsure).

We will now begin.

Getting started with AntConc

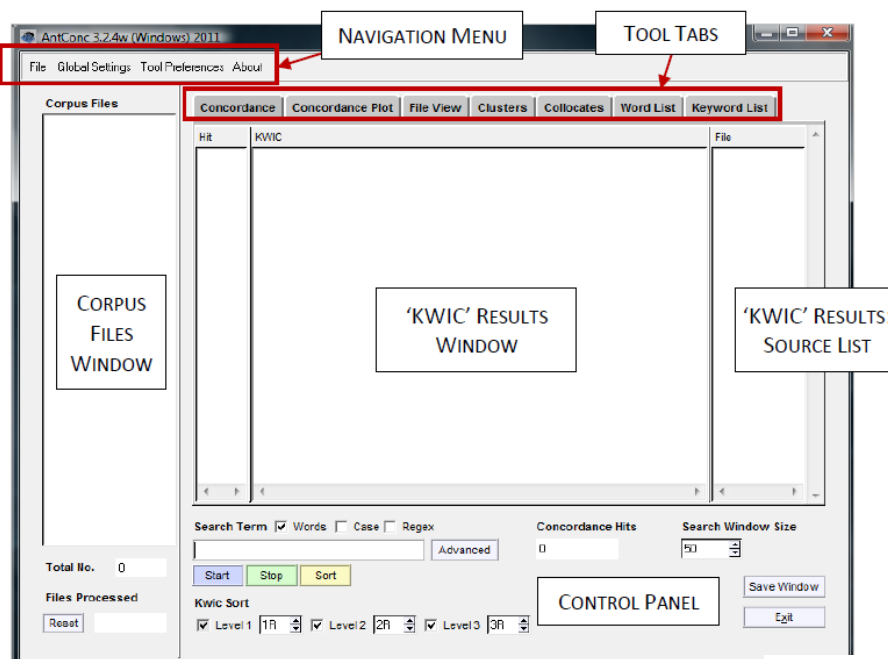
Today we are going to look at how to analyse texts using AntConc. But before we analyse a set of texts, let's open AntConc.

You can open it directly from the website or save it on to your computer or memory stick. You choose!

<http://www.antlab.sci.waseda.ac.jp/software.html>

Now follow this guide step by step.

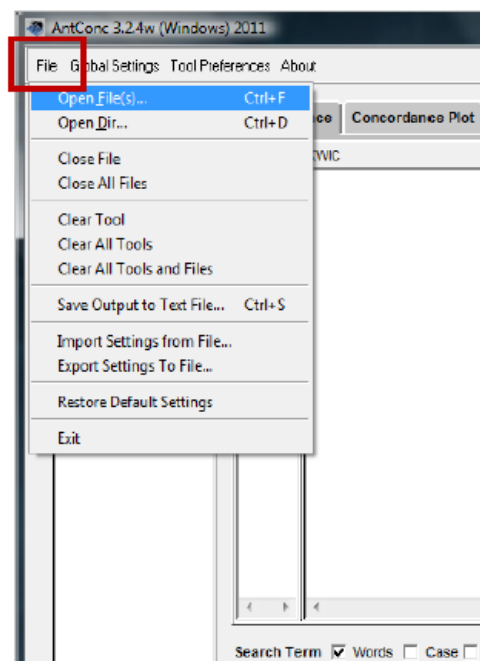
1. Open the AntConc programme. You should now have the AntConc window (see below). You can reposition, resize or maximize the window in the normal way.



Load a file

Antconcl only accepts text files (.txt) and so before you begin, make sure all of your files are saved in this format.

1. Select 'Open File(s)' from the 'File' menu.
2. Select the file
3. Click the 'Open' Files button. The chosen



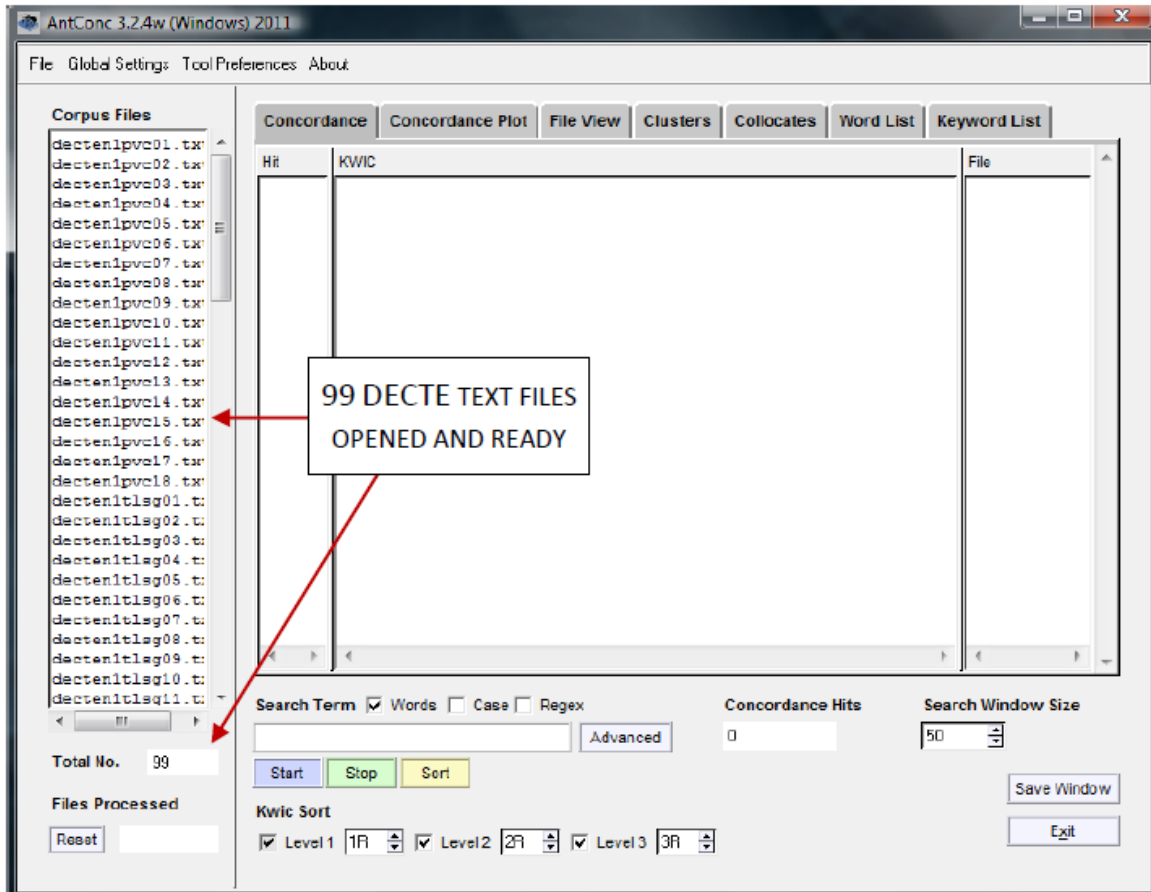
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Files are now listed in the ‘Corpus Files’

Panel on the left side of the AntConc window.

The ‘Total No’ box under the panel shows

how many files you have selected.



Concordance

To create a concordance, first make sure the ‘Concordance’ Tool Tab is selected. (The Tool Tabs are located at the top of the AntConc window. When the Concordance tool is selected, the results window and control panel will look as they do in the screenshot above.)

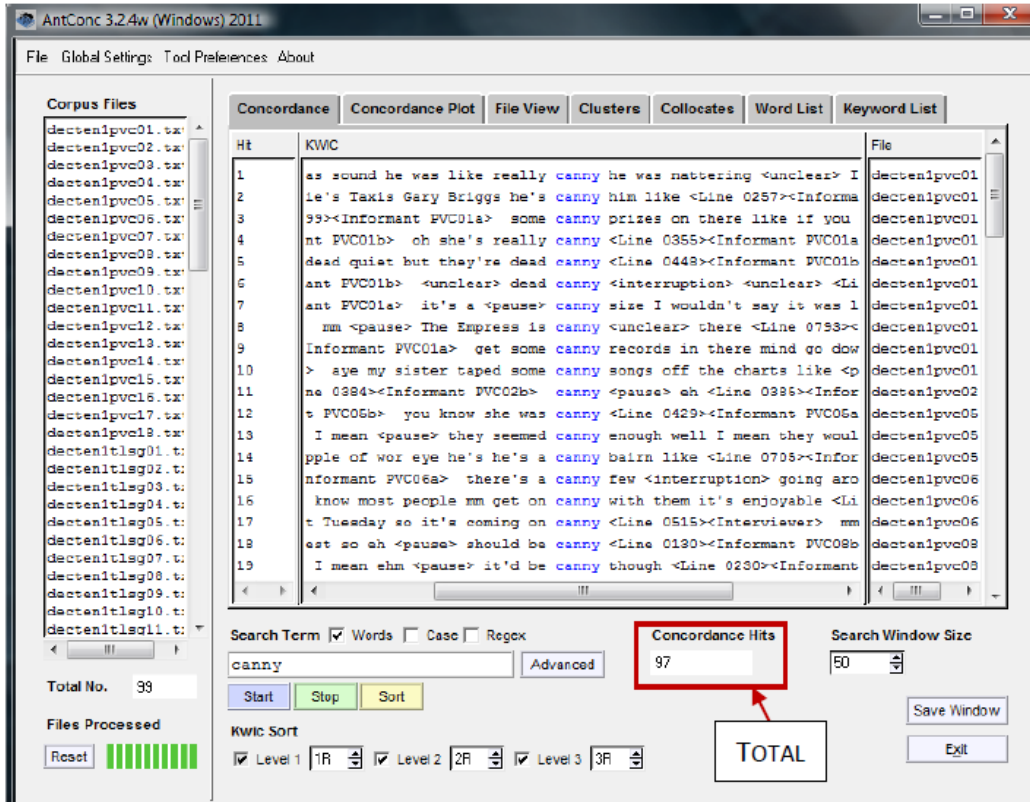
- a. To run a simple concordance using a single search term, enter the term in the search box on the left-hand side of the control panel, and click ‘Start’. For example, entering **canny** in the search box will generate a list of all occurrences of that word form in the interview text files you have chosen to search (see first screenshot on the next page).
- The results will appear as a list in the KWIC results window, showing the word in the context in which it appears in the text. The interview in which each result can be found is indicated by the file name in the ‘File’ source list window to the right of the

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

KWIC results window. The total number of results is indicated in the 'Concordance Hits' box in the control panel.

- In the KWIC results list, the key words (e.g. **canny** in the example above) appear in blue. Clicking on an instance of the key word takes you to the 'File View' for the text file in which it occurs, showing the term in its wider context (see second screenshot on the next page). To return to your concordance results list, click the 'Concordance' Tool Tab at the top of the AntConc window.

*KWIC Concordance results for **canny** (97 hits in total)*



The screenshot shows the AntConc 3.2.4w (Windows) 2011 interface. The 'Concordance' tab is active, displaying a list of results for the search term 'canny'. The results are organized into three columns: 'Hit', 'KWIC', and 'File'. The 'KWIC' column shows the word 'canny' highlighted in blue within its original context. Below the list, the 'Search Term' is 'canny', and the 'Concordance Hits' box shows '97'. A 'TOTAL' button is visible in the control panel. The 'Files Processed' section shows a progress bar and 'Total No. 99'.

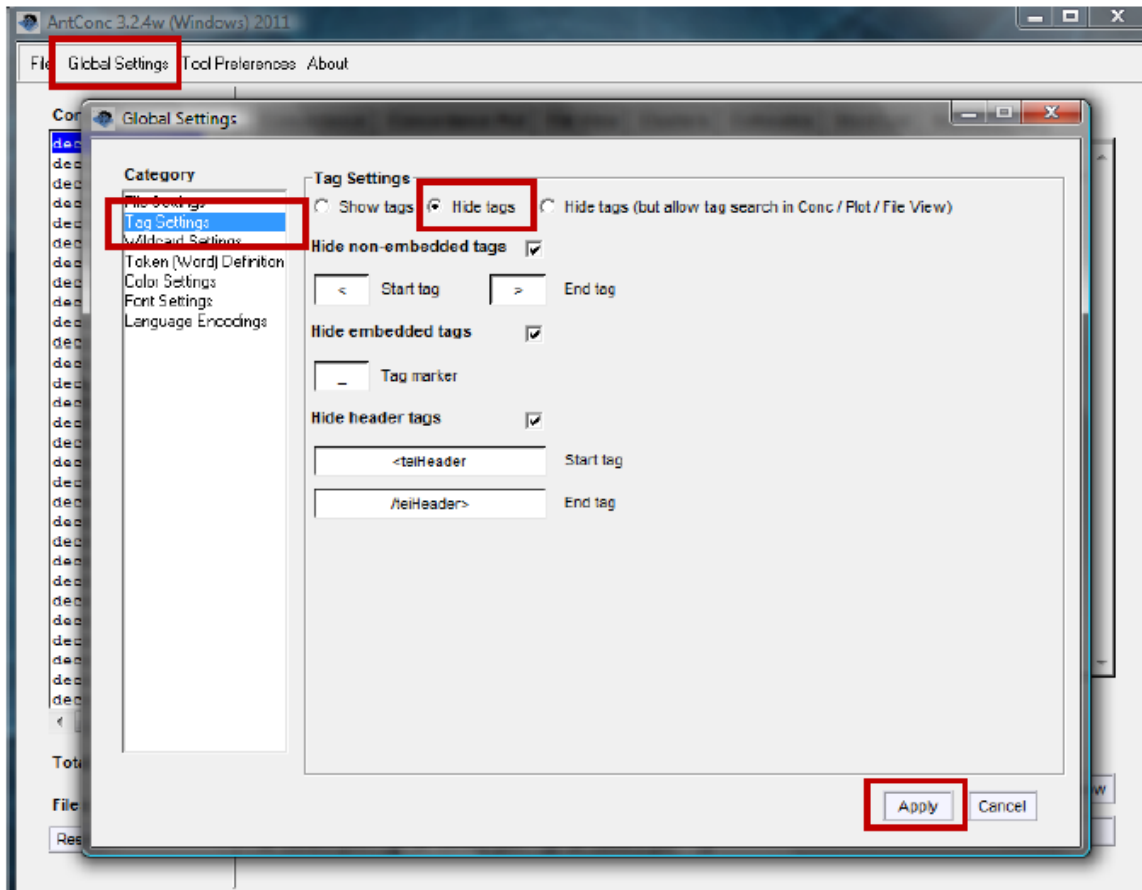
Hit	KWIC	File
1	as sound he was like really canny he was nattering <unclear> I	decten1pvc01
2	ie's Takis Gary Briggs he's canny him like <Line 0257><Informa	decten1pvc01
3	99><Informant PVC01a> some canny prizes on there like if you	decten1pvc01
4	nt PVC01b> oh she's really canny <Line 0355><Informant PVC01a	decten1pvc01
5	dead quiet but they're dead canny <Line 0448><Informant PVC01b	decten1pvc01
6	ant PVC01b> <unclear> dead canny <interruption> <unclear> <Li	decten1pvc01
7	ant PVC01a> it's a <pause> canny size I wouldn't say it was l	decten1pvc01
8	mm <pause> The Empress is canny <unclear> there <Line 0793><	decten1pvc01
9	Informant PVC01a> get some canny records in there mind go dow	decten1pvc01
10	> aye my sister taped some canny songs off the charts like <p	decten1pvc01
11	ne 0384><Informant PVC02b> canny <pause> eh <Line 0395><Infor	decten1pvc02
12	t PVC05b> you know she was canny <Line 0429><Informant PVC05a	decten1pvc05
13	I mean <pause> they seemed canny enough well I mean they woul	decten1pvc05
14	pple of wor eye he's he's a canny bairn like <Line 0705><Infor	decten1pvc05
15	nformant PVC06a> there's a canny few <interruption> going aro	decten1pvc06
16	know most people mm get on canny with them it's enjoyable <Li	decten1pvc06
17	t Tuesday so it's coming on canny <Line 0515><Interviewer> mm	decten1pvc06
18	est so eh <pause> should be canny <Line 0130><Informant PVC08b	decten1pvc08
19	I mean ehmm <pause> it'd be canny though <Line 0230><Informant	decten1pvc08

You can select 'File View' to see the wider context of the first result in the *Canny* results list.

Collocates

To generate a list of the most common collocates of a search word, first click the 'Collocates' Tool Tab at the top of AntConc window.

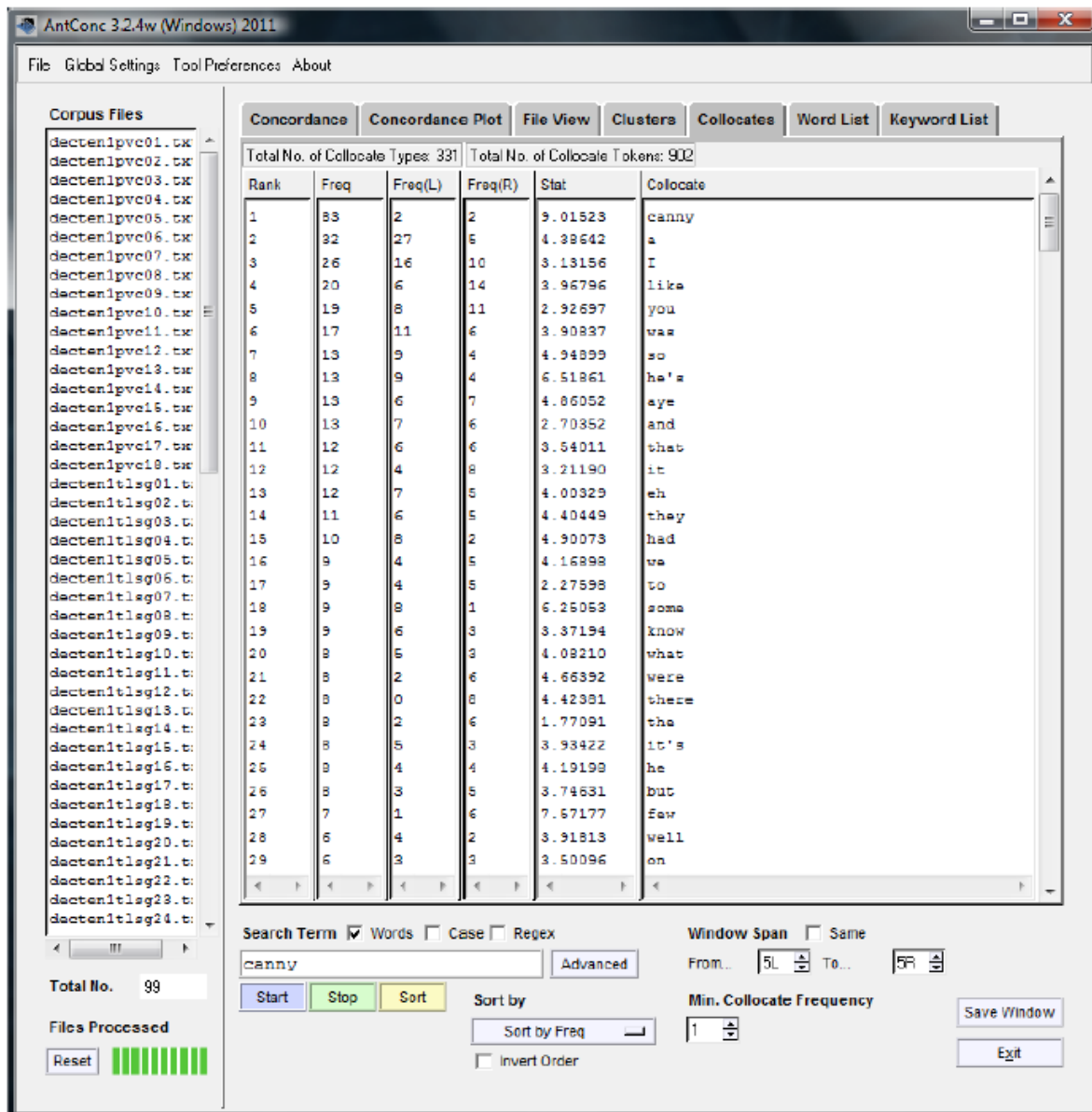
When generating a collocation list, it is a good idea to have AntConc set to ignore any tags in the text (e.g. those that identify line numbers or speakers). To do this, click 'Global Settings' in the navigation menu at the top of the AntConc window. In the resulting window, choose 'Tag Settings' from the menu on the left-hand side. By default, 'Show Tags' will be selected. Select 'Hide Tags' instead, and click 'Apply'.



Next, choose the span of words that you wish to include by altering the 'Word Span' settings in the control panel, e.g. setting 'From' to '5L', and 'To' to '5R' will search for collocates in a span of five words to the left and five to the right of the key word.

Now enter your search term in the search box, and click 'Start'. In the latest version of AntConc (3.2.4), a message will appear telling you that the program must first generate a word list. Click 'OK'. When the word list has been generated, the program will automatically return to the 'Collocates' tool and generate your collocate list.

The Collocates window lists words by the frequency with which they appear in the contexts centred on your search term (see the collocates list for **canny**, below). The list also indicates the frequency with which the collocates occur to the left or right of the key word. (By default, the 'Stat' column records a 'Mutual Information' score, which is a measure of the probability that the collocate and key word occur near each other, relative to how many times they each occur in total.)



The screenshot shows the AntConc 3.2.4w (Windows) 2011 interface. The 'Word List' tab is selected, displaying a list of collocate types and tokens for the search term 'canny'. The table below shows the top 29 results.

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	88	2	2	9.01523	canny
2	32	27	6	4.38642	a
3	26	16	10	3.13156	I
4	20	6	14	3.96796	like
5	19	8	11	2.92697	you
6	17	11	6	3.90837	was
7	13	9	4	4.94999	so
8	13	9	4	6.51861	he's
9	13	6	7	4.86052	eye
10	13	7	6	2.70352	and
11	12	6	6	3.54011	that
12	12	4	8	3.21190	it
13	12	7	5	4.00329	eh
14	11	6	5	4.40449	they
15	10	8	2	4.90073	had
16	9	4	5	4.16898	va
17	9	4	5	2.27598	to
18	9	8	1	6.26058	some
19	9	6	3	3.37194	know
20	8	6	2	4.08210	what
21	8	2	6	4.66392	were
22	8	0	8	4.42381	there
23	8	2	6	1.77091	the
24	8	5	3	3.93422	it's
25	8	4	4	4.19199	he
26	8	3	5	3.74631	but
27	7	1	6	7.67177	few
28	6	4	2	3.91813	well
29	6	3	3	3.50096	on

Create a frequency list

To create a frequency list, follow the steps:

1. Open files (see above)
2. Click the 'Word List' tab.
3. Click the 'Start' button.

Saving results

To save one of your results windows, select 'File' in the navigation menu at the top of the AntConc window. Then select 'Save Output to Text File' — OR — press CTRL+S.

So far, you have had a very brief introduction to AntConc. From what you have seen:

Discuss with a partner.

- How do you think you could use AntConc in the classroom?
- Do you think you could introduce students to it? How could students use it?