Concordancers

This Worksheet is for use with the **Concordancers** video on YouTube and Youku.

<u>Video name:</u>   **Concordancers for Academic English Study**
<u>Link to video:</u>   https://www.youtube.com/watch?v=GaQDQZLzvpE
<u>Alternate link:</u>   https://v.youku.com/v_show/id_XNTE3MDA5MDE4NA==.html

Contents

Worksheet

Video name:　　　　**Concordancers for Academic English Study**
Link to video:　　　　https://www.youtube.com/watch?v=GaQDQZLzvpE
Alternate link:　　　　https://v.youku.com/v_show/id_XNTE3MDA5MDE4NA==.html

Task 1: Before watching

Answer the following questions *before* watching the video.

i)　What resources are available to you to help you improve your academic English (e.g. dictionary, text book)? Make a list: _____

_____

_____

ii)　Study the following passage of academic writing. Some problems have been underlined. What is wrong with the words/phrases that have been underlined?

　　Different from green energy, energy from coal

　　will make many bad effects to the environment.

　　Figure 1 is showing some of these effects.

iii)　If you wanted to check the problems in ii) above, which of the resources in i) above would be most helpful?

Task 2: What is a Concordancer?　[00.00-02.12]

Watch the first part of the video (*What is a concordancer*) and complete the following notes.

Concordancer

– a **1.** _____ programme
– used to search through a **corpus**
– functions like a **2.** _____
– lists **3.** _____ containing search word
– useful to:
  ● look for **4.** _____
  ● see how **5.** _____ a word/phrase is
  ● understand how word/phrase is
     6. _____

Corpus
– a collection of **7.** _____
– plural form: **8.** _____
– many different ways to classify, e.g:
  ● **9.** _____ articles
  ● fiction
  ● non-fiction
  ● **10.** _____
  ● **11.** _____
– vary in size from few thousand
   to **12.** _____ of words

Task 3: Lextutor   [02.12-07.43]

Watch the section on the *Lextutor concordancer* and answer the following questions.

1. Which *one* of these corpora is *not* available in the Lextutor concordancer?
    - BAWE (British Academic Written English)
    - BASE (British Academic Written English)
    - COCA (Corpus of Contemporary American English)
    - COHA (Corpus of Historical American English)
    - Presidential speeches
    - Wiki Corpus

2. Complete the following table by writing down the three adjectives with the same meaning as *bad* which are common in the BAWE corpus.

Adjective	Frequency
	63
	26
	15

3. Which of the following are features of the Lextutor concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	
b) Has more than one corpus	
c) Uses academic corpus/corpora	
d) Can show frequency in different disciplines (Economics, Business, etc.)	
e) Can show frequency in different writing types (report, research paper, etc.)	
f) Is useful for collocation study (i.e. how words combine together)	

Task 4: BNC   [07.43-14.07]

1. What does BNC mean?_____
2. What type of texts are contained in the BNC corpus?_____
    _____
    _____
3. What does KWIC stand for?_____
4. What do the following areas of the BNC concordancer do?
    - List   _____
    - Chart   _____
    - Collocates   _____
    - KWIC   _____
5. Which of the areas in **6.** above would be most useful for finding out the following?
    a) Whether *difference* is more common in academic or non-academic English. _____
    b) Whether *difference* is followed by the preposition *between* or *within*.   _____
    c) Whether *major difference* is more common (in academic English) than
       significant difference.   _____
6. Which of the following are features of the BNC concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	
b) Has more than one corpus	
c) Uses academic corpus/corpora	
d) Can show frequency in different disciplines (Economics, Business, etc.)	
e) Can show frequency in different writing types (report, research paper, etc.)	
f) Is useful for collocation study (i.e. how words combine together)	

Task 5: MICUSP   [14.07-16.53]

1. What does MICUSP mean?_____

2. Which is better for understanding frequency: *raw* search results or search results *per 10,000 words*? Why?_____
_____

3. Which two disciplines does the word *effect* occur most frequently in?
- _____
- _____

4. Which two types of writing does the word *effect* occur most frequently in?
- _____
- _____

5. MICUSP has several different types of search filter. In addition to *Disciplines* (**Q3**) and *Paper types* (**Q4**), what other search filters does it have?
- _____
- _____
- _____

6. Which of the following are features of the MICUSP concordancer?

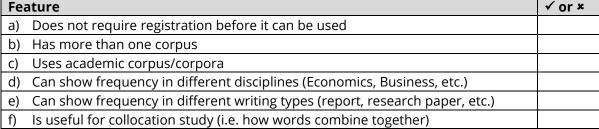| Feature	✓ or ✗
a) Does not require registration before it can be used	
b) Has more than one corpus	
c) Uses academic corpus/corpora	
d) Can show frequency in different disciplines (Economics, Business, etc.)	
e) Can show frequency in different writing types (report, research paper, etc.)	
f) Is useful for collocation study (i.e. how words combine together)	

Task 6: SKELL   [16.53-19.35]

1. What does SKELL mean?_____

2. What type of texts are contained in the SKELL corpus?_____
_____

3. How many example sentences are shown in the *Examples* section?_____

4. How do the example sentences differ from the other concordancers?_____
_____

5. What is the biggest difference between SKELL and the other concordancers?_____
_____

6. What type of information is shown on the *Word Sketch* page?_____
_____

7. Which of the following are features of the SKELL concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	
b) Has more than one corpus	
c) Uses academic corpus/corpora	
d) Can show frequency in different disciplines (Economics, Business, etc.)	
e) Can show frequency in different writing types (report, research paper, etc.)	
f) Is useful for collocation study (i.e. how words combine together)	

Task 7: Summary [19.35-21.02]

Look again at the following passage. How can it be improved?

<u>Different from</u> green energy, energy from coal will <u>make</u> many <u>bad</u> effects <u>to</u> the environment. Figure 1 <u>is showing</u> some of these effects.

Task 8: Extension tasks

Use the concordancers in this video to answer the following questions. Not all concordancers are suitable for all questions. Suitable concordancers are shown.

1. Complete the table with frequencies of collocations from the ACL (Academic Collocation List). Rank them from 1-5 according to which is most/least frequent.

Collocation	Frequency	Rank
strong correlation(s)		
strong evidence		
strong reaction(s)		
economic crisis/crises		
economic growth		

[**suitable concordancers: (1) Lextutor; (2) BNC**]

2. Complete the following sentences with a suitable preposition (*at, of, in,* etc.) for collocations from the ACL (shown in bold). [**(1) Lextutor; (2) BNC; (3) MICUSP**]
   a) There has been a **significant shift** _____ how this is perceived.
   b) A **significant portion** _____ the students have never written essays before.
   c) This may have a **significant influence** _____ how the two countries interact.
   d) There has been a **significant improvement** _____ his health condition.
   e) Mobile phones play a **significant role** _____ how we live our lives today.
   f) They have made a **significant contribution** _____ our knowledge in this area.
   g) Pronunciation is a **crucial factor** _____ how easy someone is to understand.
   h) There is a **crucial difference** _____ Einstein's and Newton's theory of gravity.
   i) This is **closely associated** _____ the level of unemployment.
   j) Economic growth is **closely tied** _____ standard of living.

3. Which discipline (out of those shown) are the words most common in? [**(1) Lextutor; (2) MICUSP**]

- **sector** (a) Economics (b) English (c) Physics
- **electric** (a) Economics (b) English (c) Physics
- **communication** (a) Economics (b) English (c) Physics
- **growth** (a) Economics (b) English (c) Physics
- **system** (a) Economics (b) English (c) Physics
- **affect** (a) Biology (b) Linguistics (c) Psychology
- **natural** (a) Biology (b) Linguistics (c) Psychology
- **climate** (a) Biology (b) Linguistics (c) Psychology
- **speech** (a) Biology (b) Linguistics (c) Psychology
- **treatment** (a) Biology (b) Linguistics (c) Psychology

Further reading/ viewing

The following are URLs for the four concordancers in the video.

Lextutor   https://www.lextutor.ca/conc/eng/
BNC   https://www.english-corpora.org/bnc/
MICUSP   https://elicorpora.info/
SKELL   https://skell.sketchengine.eu/#home

The following page gives more information on the four concordances in this video.
https://www.eapfoundation.com/news/reviews/concordancers/
The following is a review of the book mentioned at the end of the video.
https://www.eapfoundation.com/news/reviews/corpora/
The following is a sample chapter from the book mentioned at the end of the video.
https://www.lextutor.ca/TF_preview.pdf

Answers

i)   There are many possible answers here. The most likely are: *dictionary, text book, internet (incl. specific websites), mobile phone Apps, teacher, other students*. The point of this task is to make students aware that a concordancer is another language learning tool to add to others they already have.

ii) Answers vary

iii) Depends on answers to i). *Dictionary* and *teacher* may be most helpful, but may not be able to supply all answers.

Task 2: What is a Concordancer?   [00.00-02.12]

Concordancer

– a **1.** computer   programme
– used to search through a **corpus**
– functions like a **2.** search engine
– lists **3.** sentences   containing search word
– useful to:
  • look for **4.** patterns
  • see how **5.** common   a word/phrase is
  • understand how word/phrase is **6.** used

Corpus
– a collection of **7.** texts
– plural form: **8.** corpora
– many different ways to classify, e.g:
  • **9.** newspaper   articles
  • fiction
  • non-fiction
  • **10.** web pages
  • **11.** academic
– vary in size from few thousand to **12.** billions   of words
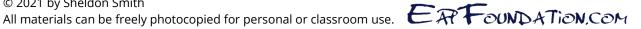
Task 3: Lextutor   [02.12-07.43]

Watch the section on the *Lextutor concordancer* and answer the following questions.

1.   Which *one* of these corpora is *not* available in the Lextutor concordancer?
  • **COHA (Corpus of Historical American English)**

2.   Complete the following table by writing down the three adjectives with the same meaning as *bad* which are common in the BAWE corpus.

Adjective	Frequency
negative	63
detrimental	26
adverse	15

3.   Which of the following are features of the Lextutor concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	✓
b) Has more than one corpus	✓
c) Uses academic corpus/corpora	✓
d) Can show frequency in different disciplines (Economics, Business, etc.)	✓
e) Can show frequency in different writing types (report, research paper, etc.)	✗
f) Is useful for collocation study (i.e. how words combine together)	✓

Task 4: BNC   [07.43-14.07]

1.   What does BNC mean? **British National Corpus**

2.   What type of texts are contained in the BNC corpus?
Spoken, fiction, magazines, newspapers, academic

3.   What does KWIC stand for? **Keyword in Context**

4.   What do the following areas of the BNC concordancer do?
- List   **shows a list of sentences with the keyword**
- Chart   **compares frequency across different sections**
- Collocates   **lists words which collocate with the keyword**
- KWIC   **similar to list but highlights parts of speech (noun in blue etc.)**

5.   Which of the areas in **6.** above would be most useful for finding out the following?

a) Whether *difference* is more common in academic or non-academic English.   **Chart**

b) Whether *difference* is followed by the preposition *between* or *within*.   **KWIC**

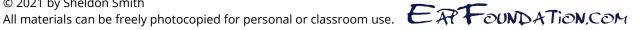c) Whether *major difference* is more common than *significant difference*.   **Collocates**

6.   Which of the following are features of the BNC concordancer?

Feature	✓ or ✗
g) Does not require registration before it can be used	✗
h) Has more than one corpus	✗
i) Uses academic corpus/corpora	✓
j) Can show frequency in different disciplines (Economics, Business, etc.)	✗
k) Can show frequency in different writing types (report, research paper, etc.)	✗
l) Is useful for collocation study (i.e. how words combine together)	✓

Task 5: MICUSP   [14.07-16.53]

1.   What does MICUSP mean? Michigan Corpus of Upper-level Student Papers

2.   Which is better for understanding frequency: *raw* results or *per 10,000*? Why?
Results *per 10,000* is better than *raw*. Some disciplines have more words than others, and *raw* results does not necessarily show how frequent a word is.

3.   Which two disciplines does the word *effect* occur most frequently in?
- **Economics (11.89 per 10,000 words)**
- **Physics (10.49 per 10,000 words)**

4.   Which two types of writing does the word *effect* occur most frequently in?
- **Research paper (39%)**
- **Report (37%)**

5.   In addition to *Disciplines* (**Q3**) and *Paper types* (**Q4**), what other filters does it have?
- **Student levels (undergraduate, 1st year graduate, etc.)**
- **Nativeness (native vs. non-native speaker)**
- **Textual features (abstract, etc.)**

6.   Which of the following are features of the MICUSP concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	✓
b) Has more than one corpus	✗
c) Uses academic corpus/corpora	✓
d) Can show frequency in different disciplines (Economics, Business, etc.)	✓
e) Can show frequency in different writing types (report, research paper, etc.)	✓
f) Is useful for collocation study (i.e. how words combine together)	✗

Task 6: SKELL   [16.53-19.35]

1. What does SKELL mean? **Sketch Engine for language learning**

2. What type of texts are contained in the SKELL corpus? **News, books, blogs, Wikipedia, other web pages, BNC**

3. How many example sentences are shown in the *Examples* section? **40**

4. How do the example sentences differ from the other concordancers? **They are shorter, rather than being extended sentences in the full context.**

5. What is the biggest difference between SKELL and the other concordancers? **SKELL does not use an academic corpus.**

6. What type of information is shown on the *Word Sketch* page? **Verbs with the keyword as subject/object, adjectives which associate with the keyword, modifiers of the keyword (e.g. adj + noun), other words which combine with it (e.g. *cause* + *effect*).**

7. Which of the following are features of the SKELL concordancer?

Feature	✓ or ✗
a) Does not require registration before it can be used	✓
b) Has more than one corpus	✗
c) Uses academic corpus/corpora	✗
d) Can show frequency in different disciplines (Economics, Business, etc.)	✗
e) Can show frequency in different writing types (report, research paper, etc.)	✗
f) Is useful for collocation study (i.e. how words combine together)	✓

Task 7: Summary   [19.35-21.02]

Look again at the following passage. How can it be improved?

Unlike   *have   adverse/detrimental/negative*

Different from green energy, energy from coal will *make* many *bad* effects

on   *shows/illustrates/depicts*

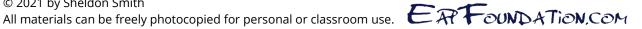*to* the environment. Figure 1 *is showing* some of these effects.

Task 8: Extension tasks

1. Complete the following table with frequencies.

Collocation	Frequency (Lextutor)	Frequency (BNC)*	Rank
strong correlation(s)	24**	15	4
strong evidence	31***	52	3
strong reaction(s)	5	4	5
economic crisis/crises	41****	94	2
economic growth	324*****	239	1

* **using +1 to the right**

** **shows 27, but 3 are not the correct form/collocation**

*** **shows 34, but 3 are not the correct form/collocation**

**** **shows 45, but 4 are not the correct form/collocation**

***** **shows 340, but 16 are not the correct form/collocation**

Note: The frequencies differ in BAWE and BNC. However, the examples have been chosen because the rank is the same in both.

2. Complete the following sentences with a suitable preposition (*at, of, in,* etc.). All of
these also use collocations from the ACL (shown in bold).
[*suitable concordancers: (1) Lextutor; (2) BNC; (3) MICUSP*]

a) There has been a **significant shift** __in__ how this is perceived.

b) A **significant portion** __of__ the students have never written essays before.

c) This may have a **significant influence** __on__ how the two countries interact.

d) There has been a **significant improvement** __in__ his health condition.

e) Mobile phones play a **significant role** __in__ how we live our lives today.

f) They have made a **significant contribution** __to__ our knowledge in this area.

g) Pronunciation is a **crucial factor** __in__ how easy someone is to understand.

h) There is a **crucial difference** __between__ Einstein's and Newton's theory of gravity.

i) This is **closely associated** __with__ the level of unemployment.

j) Economic growth is **closely tied** __to__ standard of living.

3. Which discipline (out of those shown) are the following words most common in?
[*suitable concordancers: (1) Lextutor; (2) MICUSP*]

• **sector**	**(a) Economics**	(b) English	(c) Physics
• **electric**	(a) Economics	(b) English	**(c) Physics**
• **communication**	(a) Economics	(b) English	**(c) Physics**
• **growth**	**(a) Economics**	(b) English	(c) Physics
• **system**	**(a) Economics**	(b) English	(c) Physics
• **affect**	(a) Biology	(b) Linguistics	**(c) Psychology**
• **natural**	**(a) Biology**	(b) Linguistics	(c) Psychology
• **climate**	**(a) Biology**	(b) Linguistics	(c) Psychology
• **speech**	(a) Biology	**(b) Linguistics**	(c) Psychology
• **treatment**	(a) Biology	(b) Linguistics	**(c) Psychology**

Teacher's Notes

Video name: **Concordancers for Academic English Study**
Link to video: https://www.youtube.com/watch?v=GaQDQZLzvpE
Alternate link: https://v.youku.com/v_show/id_XNTE3MDA5MDE4NA==.html
Video length: 22.57
Learning Outcomes: (1) Understand what a concordancer (and a corpus) is.
                     (2) Be aware of features of different concordancers.
                     (3) Be able to use concordancers to study academic vocabulary.

Suggested sequence/timings (for teachers)
This is a long video. It would be suitable for more than one lesson, and this approach is recommended. The following would be a possible sequence.

    Lesson 1: Introduction to corpora and Lextutor concordancer + *vocabulary study*
    Lesson 2: Introduction to BNC concordancer + *vocabulary study*
    Lesson 3: Introduction to MICUSP and SKELL concordancers + *vocabulary study*

Lesson 1
1. Warm-up (10')
a) Do *Task 1: Before watching*. This task can be done before giving the handout to students, e.g. copy text/questions to the board or PPT. The aim of this task is to get student to reflect on the tools they have available to study and improve their English, before presenting one they are unlikely to have encountered before.
b) Get some feedback from students. Note: it is not necessary (nor desirable) to go through language answers, as these are presented in the video. It is also not important if students know the answers already or not.

2. Understanding corpora (10')
Video segment: 00.00-02.12
a) Direct students to watch the video and take notes.
b) Give out the handout and have students (in pairs) complete Task 2.
c) If necessary, allow students to watch a second time to check answers.
d) Go through answers with students.
Note: It is useful to do this first, so that students have a clear understanding of what a concordancer (and a corpus) is.
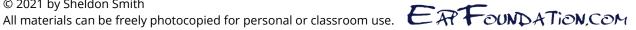
3. Lextutor concordancer (20')
Video segment: 02.12-07.43
a) Direct students to watch the video and take notes.
b) Students complete Task 3.
c) Go through answers with students.

4. Using a concordancer (20')
a) Move on to Task 8. Have students complete this using Lextutor.
b) Feedback.
Note: it might also be helpful to go over Task 7 first, if time.

Lesson 2
1. Review (10')
Review elements of previous lesson (e.g. what is a concordancer, what is a corpus).

2. BNC concordancer (20')
Video segment: 07.43-14.07
a) Direct students to watch the video and take notes.
b) Students complete Task 4.
c) Go through answers with students.

3. Extension activities (20')
There are no further activities. The teacher would need to devise more for use with BNC.

Lesson 3
1. Review (10')
Review elements of previous lessons (e.g. how are Lextutor and BNC concordancers similar/different).

2. MICUSP concordancer (10')
Video segment: 14.07-16.53
a) Direct students to watch the video and take notes.
b) Students complete Task 5.
c) Go through answers with students.

3. SKELL concordancer (10')
Video segment: 16.53-19.35
a) Direct students to watch the video and take notes.
b) Students complete Task 6.
c) Go through answers with students.

4. Extension activities (20')
There are no further activities. The teacher would need to devise more for use with BNC.