# Pedagogical Applications of Computer-Assisted Textual analysis using CLic, Voyant, and Antconc

Najla Jarkas

Mellon Fellow for the digital liberal arts programs

The amical consortium

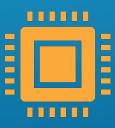
November 30, 2022



#### overview

- Recap
- Samples of assignments and rubrics
- How do we prepare students for computer-assisted textual analysis assignments?
- How do we support students with interpreting and reporting their results?

#### REcap



What is computer-assisted textual analysis?



How do we compile a corpus for analysis?



Quick revision of tools: CLiC; Voyant Tools; Antconc

#### WHAT IS COMPUTER-ASSISTED TEXTUAL ANALYSIS?

- Computer-Assisted textual analysis means using computers to "read" sets of digitized texts for us.
- It uncovers word trends, recurring word patterns and collocates in singular or multiple texts across centuries and boundaries
- It utilizes visualizations such as graphs, maps, wordclouds, scatter plots etc to represent lexical patterns
- It applies quantitative and qualitative approaches to texts
- It facilitates, interogates, and compliments the reading of large sets of texts, but it does not replace human reading

#### Why do we need it?

"Computers are good at mechanical processes, such as searching, counting, and sorting into alphabetical or numerical order. They preform these tasks not only much faster than a human being, but also very much more accurately." (Hockey, 200, p.3-4)

#### HOW DO WE COMPILE A CORPUS FOR ANALYSIS?

- Compiling a corpus relies very much on your research question(s) and scope of enquiry
  - Save files as plain texts in ways
    - with author's surname \_ short title of work\_ pub year (Antconc): Dickens\_Taleoftwocities\_1859
    - with pub year\_author's surname \_ short title (Voyant): 1859\_Dickens\_Taleoftwocities
  - Save files in two folders/corpora to compare perspectives:
    - Corpus/Folder A: texts representing one perspective/hypothesis
    - Corpus/Folder B: texts representing the opposite perspective
  - To compare keywords:
    - Corpus/Folder C: texts for analysis of special keywords
    - A Reference Corpus/Folder D: texts for comparison
  - More options will appear as you revise your research questions and develop your analysis

#### How do we start?

- Start small and work up to larger data sets
  - Text analysis projects are iterative. You should always start small and work up to larger and larger data sets. The data set you curate we will call a corpus. It could be one or more texts unified by some criterion relevant to your research. You may treat one text as many or many as one. You may slice it in different ways. The corpus selected, captured, and generated will constrain and enable questions so you should have some idea of what you want to study when you start gathering a corpus. As you ask questions you will realize you need to add or remove from the corpus. It will change as your thinking does. You are in a conversation with text." (Geoffrey Rockwell and Stéfan Sinclair. Dialogica)

#### Assignments using textual analysis tools

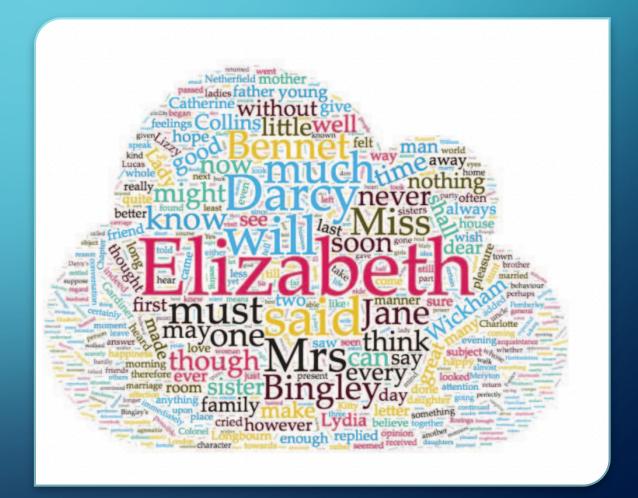
- Don't feel obliged you have to include a digital assignment: "Digital work for digital's sake is never a good idea and adding a tool to a process in which it is unnecessary can make coursework cumbersome and tedious." <a href="https://mcgrawect.princeton.edu/digital-assignment-auides/">https://mcgrawect.princeton.edu/digital-assignment-auides/</a>
- Reach out to digital librarians and IT staff for support in and out of the classroom
- Start with learning outcome(s) that you want the assignment to address
- "assignments can encourage collaborative work, exploring multiple literacies, writing for public audiences, and learning to effectively present critical, creative, and community-engaged scholarship. These non-technical outcomes should form the basis of and motivation for the development of digital assignments." <a href="https://guides.library.ucsc.edu/DS/Resources/GettingStarted">https://guides.library.ucsc.edu/DS/Resources/GettingStarted</a>
  - What specific perspective(s) do you want your students to gain on contents or materials in your course?
  - Figure 1 Textual analysis helps student to see patterns, word trends, uncover themes, etc.
  - Which of the tools introduced in the previous workshop can match what you expect your students to learn?
  - Think of the learning curve
  - How much time do you want the assignment to take for students to complete including the time technical skills will be acquired?
  - Do you want it to be an indiviual or group assignment?
  - Think about grade weights and scale of the assingments
  - What do you want the end product to look like? Digital assignments can be multimodal. Where do you want students to create and post their textual analysisexplorations and results?
  - How will you integrate and grade the critical reflections students have about the process and digital tools

#### CLiC

- Discussion
  - ► What can you do with CLiC?
  - ►What can't you do?
  - When and where will you be using it?

#### **CLiC** assignments

- Teaching and doing research with CLiC:
  - Explore themes, setings, characterization in one novel: A Christmas Carol by Dickens
    - CLiC in the Classroom: https://blog.bham.ac.uk/clic-dickens/2018/03/05/clic-in-the-classroom/
    - Assignment handout: <a href="https://blog.bham.ac.uk/clic-dickens/wp-content/uploads/sites/44/2018/03/CLiC">https://blog.bham.ac.uk/clic-dickens/wp-content/uploads/sites/44/2018/03/CLiC</a>
      Dickens-Project-Christmas-Carol.pdf
  - Activity:
    - Keywords are useful as a first step to get an overview of a text and what it is about.
    - To get an understanding of the concept of keywords, word clouds are useful to start with. You might want to ask your students to produce a word cloud for a book of their choice and then the whole class can see whether they will know the book from the word cloud. The word cloud in Figure 15 is rather easy to recognise as a representation of *Pride and Prejudice*.



#### CLiC assignments

Literature assignment: "for example, you might notice a feature, theme, issue or pattern in your reading of a novel, but your sense of it might be rather vague and you would like to explore it further: CLiC can help you do this. It can also help you to refine your thinking, find evidence to support your intuitions, and reveal further patterns in the text that you might not have noticed explicitly."

#### Activity:

- Start again by going to the CLiC Concordance tab (http://clic.bham.ac.uk/concordance).
- In "Search the corpora", select "David Copperfield".
- In "Only in subsets", select "All text".
- In "Search for terms", insert Dick
- In "Filter rows", try out a range of terms such as head, face, eyes, mouth, or looked, watched, seemed. Hit Return after each one.
- > Try other synonyms or alternatives as they occur to you.
- Search for "chin". How frequently does it appear in the Dickens corpus
- Then search for "chin" in the 19th Century Reference Corpus and make a note of the frequencies
- Now search for same word in Austen's corpus
- Come up with observations about gender roles and body parts such as chins in three corpora.
  - What adjectives and verbs are associated with chins
  - Why do the chins that are mentioned in the Dickens novels often belong to male characters and tend to be evaluated less harshly than female chins?

#### CLiC assignments

- So what? Well, this is where your own interpretative sense of the worlds of Dickens and Austen respectively comes into play.
- Why do you think there are more chins in Dickens than elsewhere?
- What does that seem to point to about how Dickens creates characterisation?
- Which term might we alternatively search for, using the same technique as here, that we intuitively feel might be overly used by Jane Austen, comparatively?
- For example, students may pick up on gender role differences. The chins that are mentioned in the Dickens novels often belong to male characters and tend to be evaluated less harshly than female chins.
- In Dickens, chins are generally impolite often (as double chins; see our example on concordancing and KWIC Grouping chin in Part I, Section 4) used to signify over-indulgent characters, or as an index of wealth and higher, pampered class. Chins are often stroked, rubbed, touched, and held as a sign of internal thought, careful perusal, or scheming.
- ➤ Think of other body parts that you might search for in a similar way. This can either be ones that you just think of randomly, or you might feel that your reading of a particular novel suggests a focus on a body part or a gesture that feels significant
- https://www.birmingham.ac.uk/documents/college-artslaw/elal/clic/clic-activity-book-license.pdf

#### Voyant tools

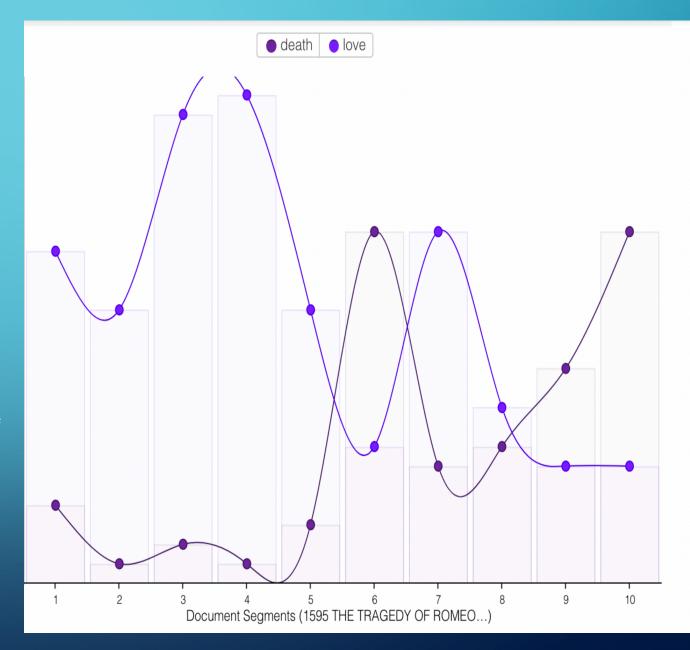
- Discussion
  - What can you do with Voyant Tools?
  - ►What can't you do?
  - When and where will you be using it?

#### Voyant assignments

#### Teaching with Voyant:

https://digitalfellows.commons.gc.cuny.edu/2018/11/01/using-voyant-tools-to-formulate-research-questions-for-textual-data/

- Love and death in "trends" Romeo and Juliet
- Compare the frequencies between "Love" and "Death". Do you notice anything interesting? They show an inverse relationship: where "Love" occurs most often, "Death" has a low frequency; and where death is prominent, love diminishes. This inverse relationship—which might seem unremarkable on the surface—indicates that there is more work to be done on examining the associations between love and death. Most importantly, it points to areas of tension in the text that the reader can re-read more closely. . These collisions between "Love" and "Death" present potential sites of further questioning: what is the exact nature of this relationship? How might the two seemingly opposite forces be connected?"
- Ask students to specluate about other "concealed" aspects about the texts to open the text to new questions and insights.



#### Voyant assignments

- Blog assignment
- https://njarkasblog.wordpress.com/engl-256b-blog-2-seeing-through-your-literary-texts/
- Visualizations from students analysis are integrated into the final project on TimelineJS

#### Antconc

- Discussion
  - What can you do with Antconc?
  - ► What can't you do?
  - When and where will you be using it?

#### Antconc assignments

- Orientalist fan fiction corpus from Gutenberg project
- Students selected genres:
  - Translations of folktales into English
  - Travel literature
  - Military reports and expeditions
  - Religious texts
  - Diaries
  - Novels, short stories, and poems
- Explored terms, phrases, and themes that they would expect to find in Orientalist fiction
- Explored new themes to uncover new approaches to the texts

Where in your sylabus does it make sense to add an assignment?

Discussion

When in your sylabus does it make sense to offer training to your students on the tools?

Discussion

### How do we prepare students for computer-assisted textual analysis assignments?

- Go through a short survey of distant reading practices. Feel free to use my slides
- Show how it could be relevant to their studies and research practices in other courses as well
- Remind them of copyright good practices
- Offer them a plain text and ask them to exlore the tools with guided questions
- Ask them to work in teams to compile their own corpus around a research question, write the process, and findings with evidence exported from each tool:
- Engage them in questions like:
  - What did you expect the tool to uncover in the text(s)
  - What new findings emerged as you explored the tools
- Which of the tools do you prefer to use? Why?

### How do we support students with interpreting and reporting their results?

- Students need to make sure they are accruately reporting their process and findings.
- They could report their findings in an inforgraphic poster, blog or online analysis essay

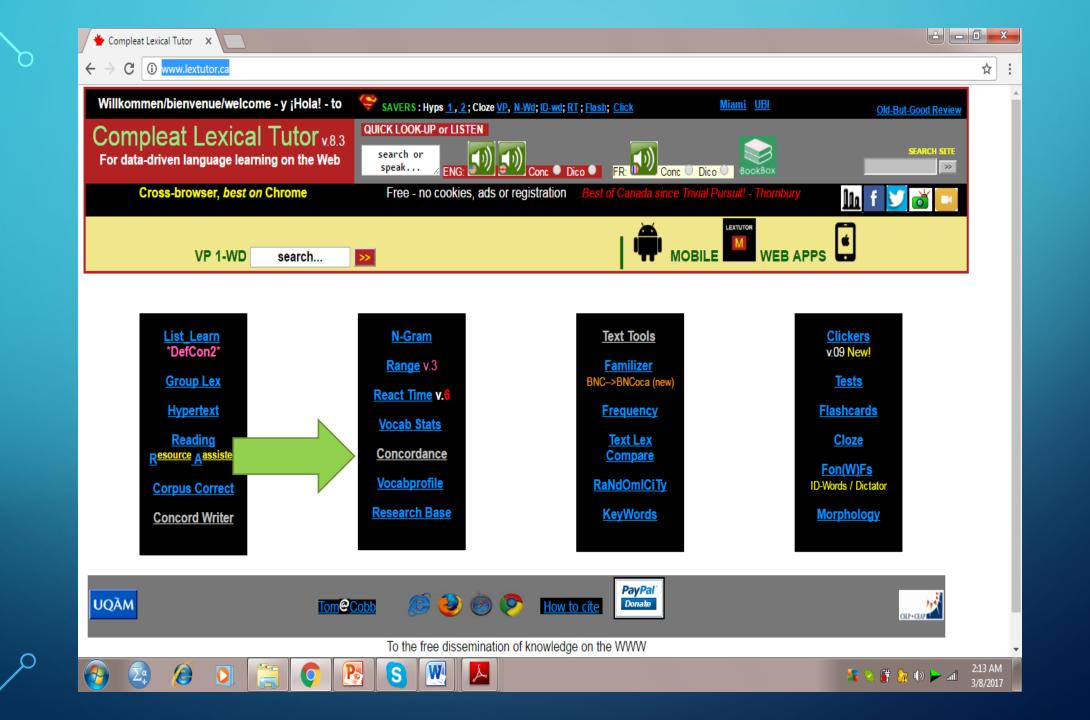
## TEXTUAL ANALYSIS USING COMPLEAT LEXTUTOR

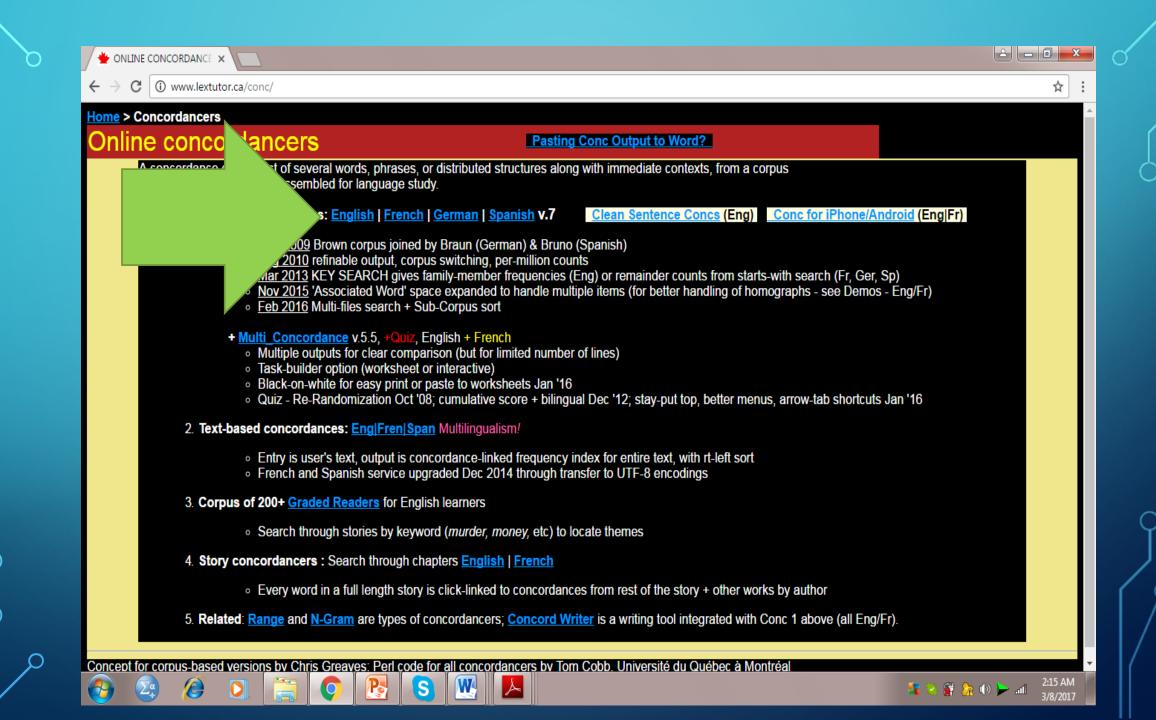
MAHEEN ZIA FCCU

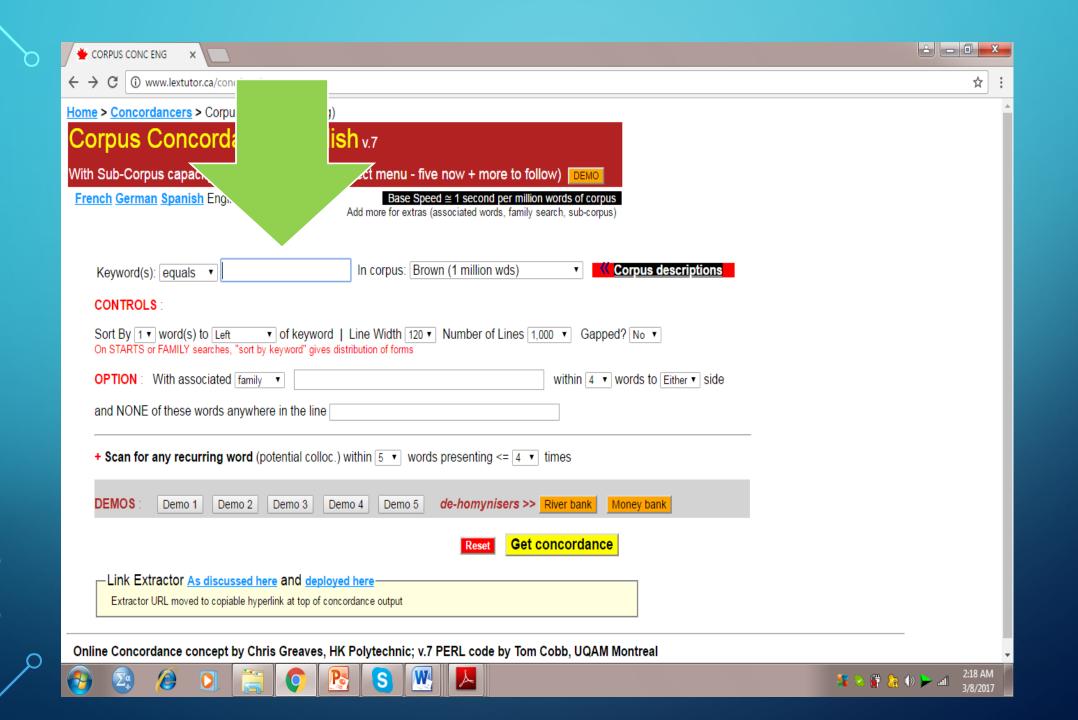
COURSE: ENGLISH WRITING COMPOSITION

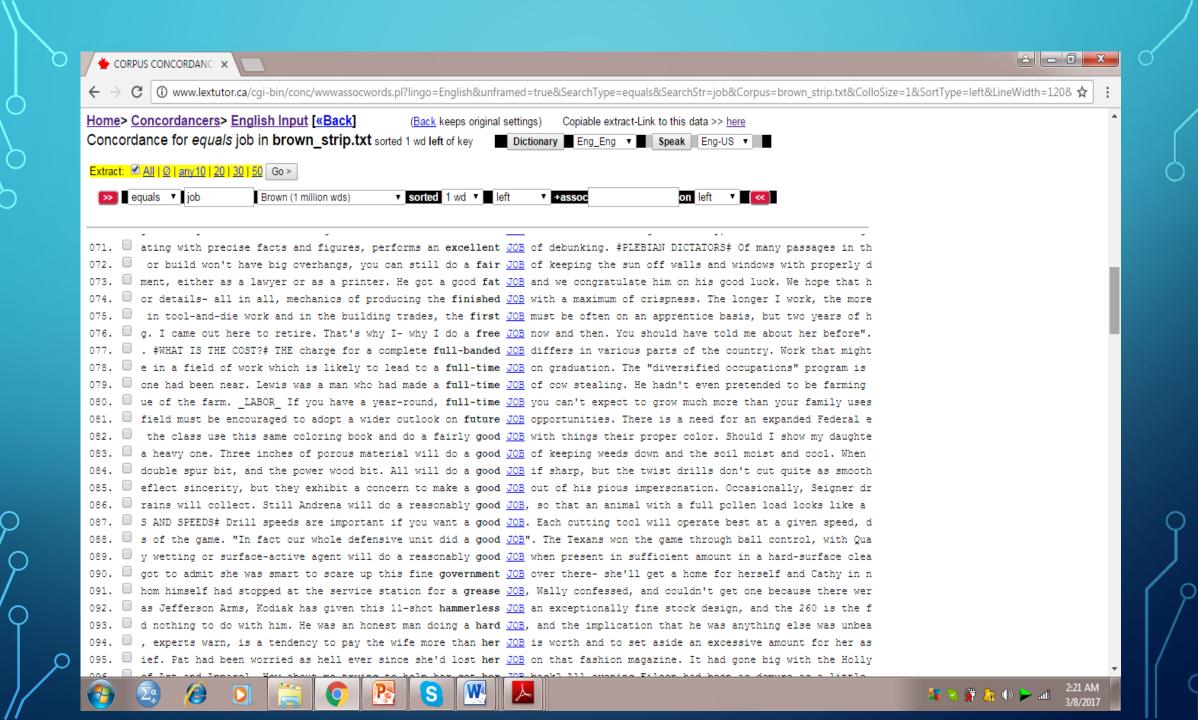
## TEACHING COLLOCATIONS THROUGH AN ONLINE CONCORDANCE TOOL CALLED COMPLEAT LEXTUTOR

http://www.lextutor.ca/

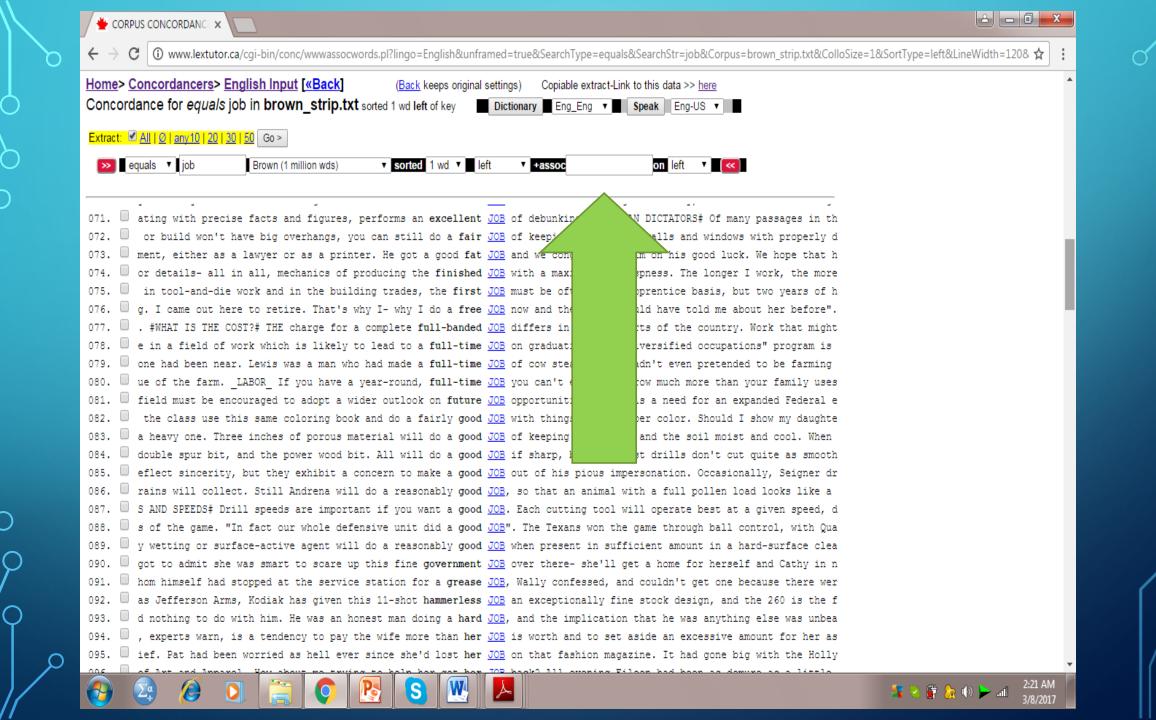


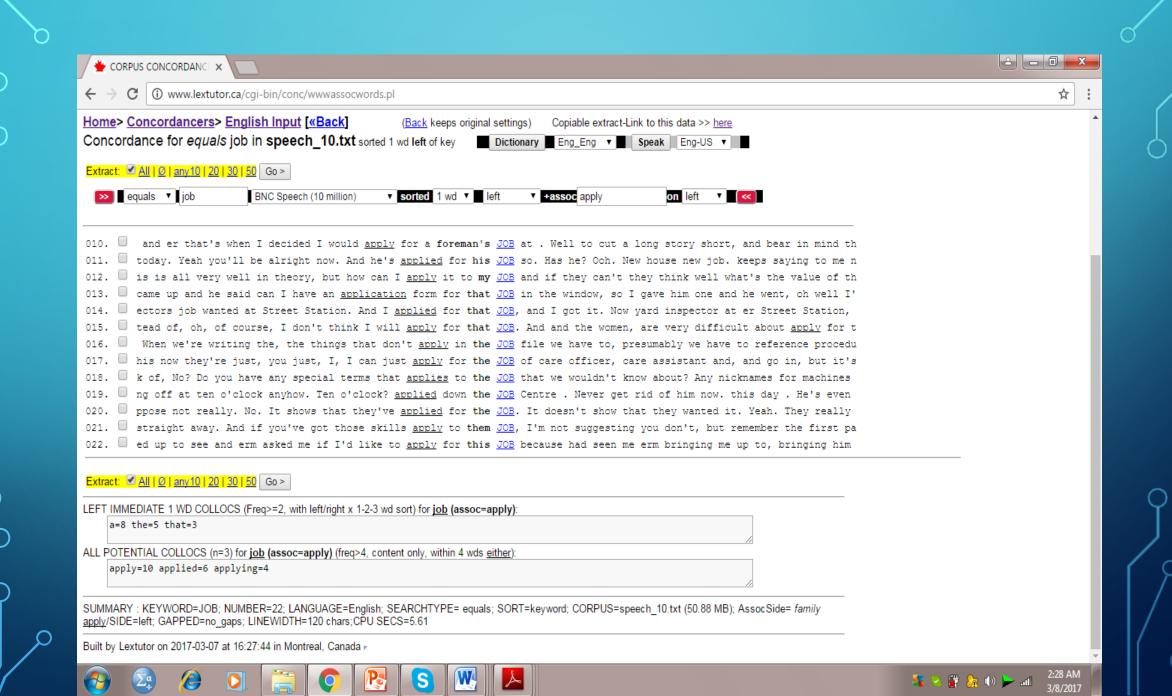












## GRID COMPLETION ACTIVITY VERB + NOUN

	Respect	Money	Support
Deserve	<b>✓</b>		
Make	*		
Have	<b>✓</b>		

#### **ACTIVITY QUESTIONS**

- Step 1: Fill the grid by putting a tick or cross mark in the spaces, using lextutor tool as an aid.
- Step 2: Run a search for the tick-marked collocations and use them in your sentences.

#### STEP 1

ales all politicians were supposed to receive credit **new** RESPECT for turning more liberal But this problem was quickl i builders who worked when art was a living thing <u>had</u> no <u>RESPECT</u> for the work of builders who went before them but pu le i know really i i liked Tom Landry you know i <u>have</u> no <u>RESPECT</u> for the for the new coaching staff since they ousted out because Shaun is boring er put er Shaun she has no RESPECT for she says because he hasn t got a job He won t ge Buit um and i really felt that by the end of it i <u>had</u> no <u>RESPECT</u> whatsoever for our jury system uh huh oh really what ng why things like this happen why some children <u>have</u> no <u>RESPECT</u> for life not even their own She wondered what would she snapped adding icy emphasis to his name You have no RESPECT for me I can see that Hardly surprising he murmured because Shaun is boring er Don put her Shaun she <u>has</u> no <u>RESPECT</u> for she says because he <u>hasn</u> t got a job he won t ge to buy their drugs with so sure sure uh and they have no RESPECT or for for any other people anyway that s true but w yot crabby about people dropping litter You kids <u>have</u> no <u>RESPECT</u> for the environment she said She made Marcus Anderso a degree of placidity which I can neither comprehend **nor RESPECT** There were misunderstandings between them Emma he sa persuade shows rather explicitly that the pope does not RESPECT the freedom of others The basic religious doctrine h ı a likeness almost before it was possible She could **not RESPECT** his eye but his love and his complaisance were unexc nich she was beginning to find irritating the obsequious RESPECT of the victim for his oppressor Sometimes appalled b the quote unquote traditional route and i have a lot of RESPECT for her because she made it as a real choice um 289. through three classes and i <u>have</u> a great deal of <u>RESPECT</u> for it uh huh uh huh um i <u>haven</u> t done much word pro we argument be used to base authority on the attitude of  $ext{RESPECT}$  An obvious objection is that the attitude of respect ould revolve around their needs and demands What kind of RESPECT for our common culture is that For old fashioned hon vn world it it s a i i have more respect i have a lot of RESPECT for MIT master s and doctorate uh yeah it s a top ra to give it every bad name in the world Some portion of RESPECT for herself however in spite of all these demerits s

#### STEP 2

Sentence:

Have respect: I have no RESPECT for her because she completely gave up.

#### TEXTUAL ANALYSIS USING ANTCONC

Course: Corpus Linguistics

#### **PROJECT**

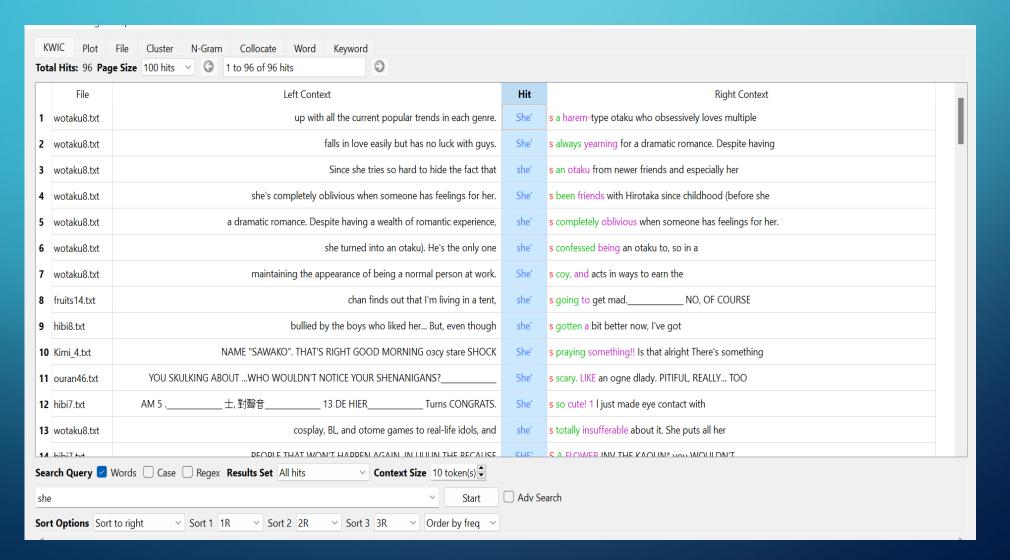
- Japanese comic books
- Type 1: targets teenage male readers
  - Type 2: targets teenage female readers
- Corpus compiled and run on AntConc

#### **PROJECT**

Analysis: testing the themes of masculinity and femininity

- 1. frequently occurring content words
- 2. collocations through the concordance tool for gender pronouns (he, she, they)

#### **SNIPPETS**





#### sources https://blog.bham.ac.uk/clic-dickens/2018/03/05/clic-in-the-classroom/ https://blog.bham.ac.uk/clic-dickens/category/teaching/ https://www.birmingham.ac.uk/documents/college-artslaw/elal/clic/clic-activity-book-license.pdf https://digitalfellows.commons.gc.cuny.edu/2018/11/01/using-voyant-tools-to-formulate-research-questions-for-textual-data/ https://mcgrawect.princeton.edu/digital-assignment-guides/ http://pedagogy-toolkit.org/tools/VoyantTools.html https://guides.libraries.psu.edu/c.php?g=1123281&p=8275191 https://guides.library.ucsc.edu/DS/Resources/Voyant#s-lg-box-wrapper-29088756 https://guides.library.ucsc.edu/DS/Resources/GettingStarted https://njarkasblog.wordpress.com/engl-256b-blog-2-seeing-through-your-literary-texts/ https://library.villanova.edu/application/files/9815/8739/5480/VoyantToolsParadiseLostTutorial.pdf