Activity Guide
Text Analysis and Visualization with Voyant Tools & AntConc workshop

Set-Up Steps:

1. Get Data: go.ncsu.edu/voyantfall2018
2. Get Voyant Server: download the VoyantServer 2.0 zip archive
3. Load data into Voyant: Click Upload button, select **fairy-tales-corpus.zip**

Activity 1: Create a custom stop-words list

Focus your attention on the word cloud tool (Cirrus). Here you see the most frequent words in the fairy tales displayed. The most common words are the largest words in the cloud.

Common stop words in English have already been removed, such as **the, an, and, or, but**, etc.

As we see by the size of the words, the most frequent terms in this corpus are
- **said**
- **little**
- **came**
- **beautiful**
- **prince**

We can also see this in the lower left section of the page:

Most **frequent words** in the corpus: said (201); little (152); came (89); beautiful (81); prince (74)

Some of these words should be added to the stop words list so we can filter them out. Which would you choose?

Instructions for Adding New Stopwords

1. Go to this document and copy all of the words in the list. We will add these words to the Stop Words list in Voyant.

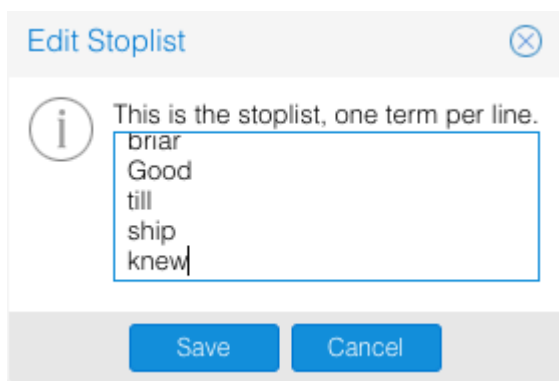2. Hover over the gray bar at the top of the **trends window** until a menu of icons appears.

3. Click on this icon:

4. Click on the "Edit List" button to the right of "Auto-detect" dropdown menu


5. In the Edit Stoplist window that pops up, copy the words you pasted from the Stopwords document , and Save:


6. Click frequencies: Raw then confirm.

The word cloud should update with those terms filtered out:

You will need to click Reset below the **Trends** window to see your updated terms in the graph:

Activity 2: Examine frequent word trends by document

1. Look at the Trends section (upper right section) for a graph of word frequencies over the corpus of fairy tales. Click "Display" at the bottom of the Trends section and choose "columns" to change the chart type.
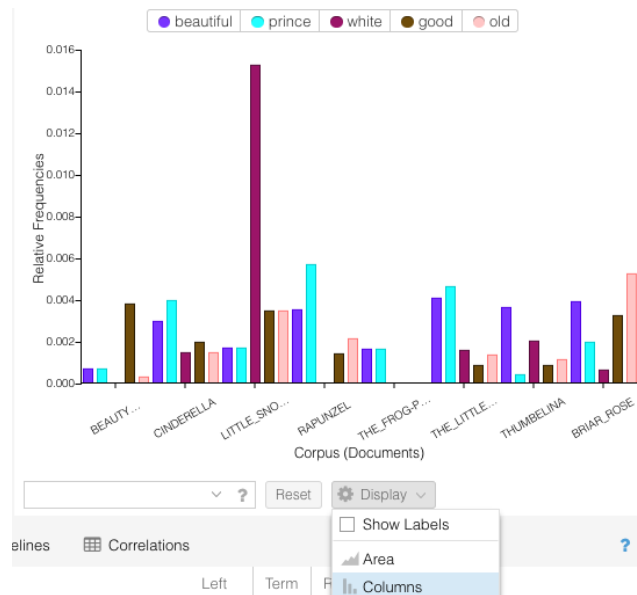
2. To select a different word to examine in Trends, click on that word in the word cloud tool.

3. Click the word "prince" in the word cloud.
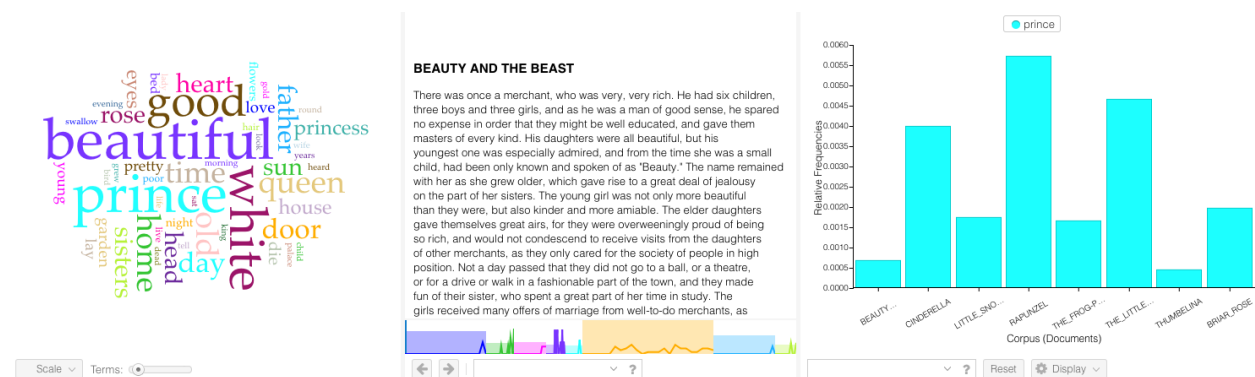
In the middle section, the **Reader**, the bars show the relative size of each text in the corpus, with a line graph showing the distribution of the word "prince" in each text.

4. To select multiple words to examine, type a word in the entry box in the **Trends** section. You can type multiple words here. (Press Enter after each word.)
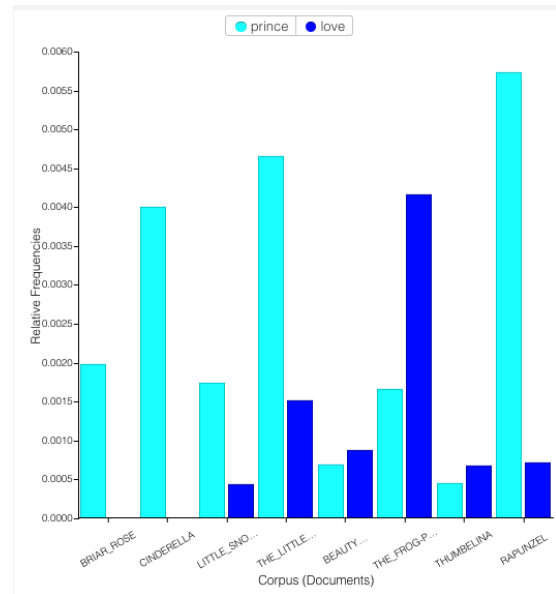
Here's an example graph related to the terms **prince** and **love**.

We can see that sometimes, "prince" appears without any mention of "love" at all.

5. Take a few moments to examine the trends of words that interest you in the word cloud.

Activity 3: Examining words in context

Term frequencies and trends can only take us so far with analysis. **Context is also important.**

Let's examine the term "love".

1. Click on "love" in the Word Cloud

2. In the **Contexts** section (lower right), type "love" into the search box. The word will appear with surrounding words for context, organized by fairy tale.

3. You can see the word through the entire corpus or select specific documents. To see "love" in all documents, click on the Scale button. Select Corpus (should be the default).

Sample Result:

	Contexts	Bubblelines	Correlations		?
	Document ↑	Left	Term	Right	
⊞	1) BEAUTY…	in so doing, for by my de…	love	for him." "No, my sister," said the thre…	
⊞	1) BEAUTY…	who told him that there h…	love	with her sisters. She begged her fat…	
⊞	1) BEAUTY…	satisfied with that.' "I sup…	love	you very much. However, I am but t…	
⊞	1) BEAUTY…	eldest had married a you…	love	with his own face, that he could thin…	
⊞	1) BEAUTY…	happy; it is amiability of c…	love	him, but I respect him, and I feel bot…	
⊞	3) LITTLE_…	And the prince answered…	love	you better than all the world; come …	
⊞	4) RAPUN…	she would take him for he…	love	me more than old Dame Gothel doe…	
⊞	5) THE_FR…	into the spring.' The frog …	love	me, and let me live with you and eat…	
⊞	5) THE_FR…	and a little voice cried ou…	love	here! And mind the words that thou …	

love ✕   ⌄ ?   29 context ⊙—— expand ⊙——   Scale ⌄

4. To see more context, click on the + sign to expand

5. Use the "context" slider bar to provide more Left and Right words in the table.
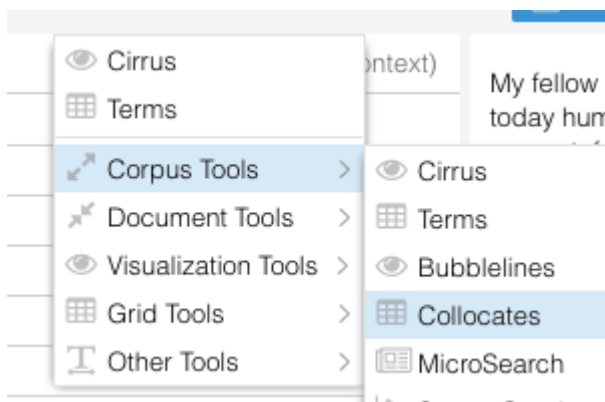
6. What can you observe from looking at the Contexts tool? Is there anything surprising or unexpected?

Activity 5: Replace Word Cloud Panel with Collocates Tool

1. Click on the little window icon in the gray strip at the top of the Cirrus word cloud tool

2. Choose "Corpus Tools" and "Collocates" to swap Cirrus with the Collocates tool

3. Type in the word "old" in the empty search window

4. Examine the collocated terms with the word "old"

5. Select "old woman."

6. Notice that the trend graph updates with collocation frequencies.

You can select more than one collocation; for example, select the top 3 (old-woman, old-lady, and old-grandmother).
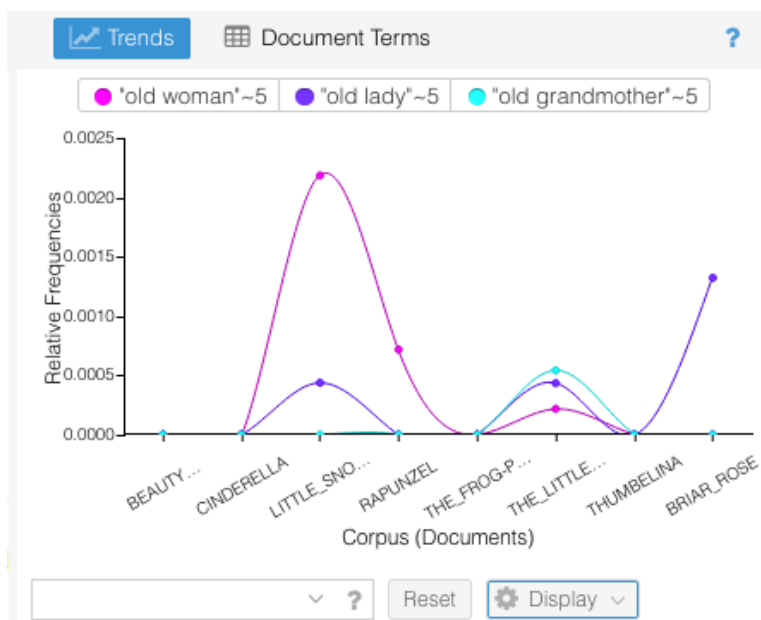
Use the context slider to change the number of words to consider on each side of the keyword (such as 6 to the left and 6 to the right). The default is 5.

7. Spend a few minutes finding collocated words for a different term.

Activity 7: Export and save your stuff!

It's possible to export images, data, URLs and bibliographic references of your project. Here are two examples:
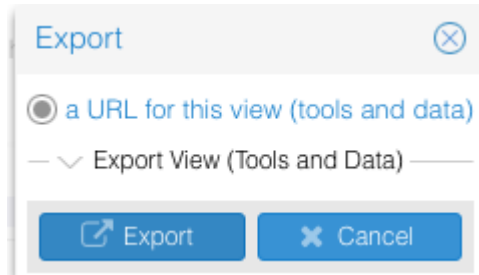
1. To Export a URL of your project (so that others can see it):
To export your project: go to the blue strip at the top of the screen:

Click the icon with the arrow coming out of it

Click Export button -- this will generate a link you can share with others

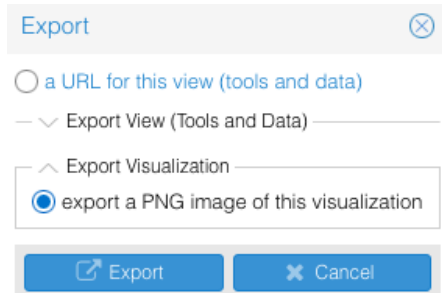2. To save a PNG file of a specific graph
Click on this button at the top of the graph:
Select Export Visualization
Select "export a PNG image of this   visualization"

NEXT TOOL: AntConc: Tool for Text Analysis

At the end of the workshop we'll look at AntConc, an open source tool for text analysis that is well suited for concordance. You can download it and follow along, or watch the demo and download it later.

Download AntConc: http://www.laurenceanthony.net/software/antconc/

Helpful list of stopwords: Buckley-Salton Stoplist