**Professional Writing Text Comparison for Word Choice, Style, and Audience**

Students will use Voyant to compare professionally written journalism across genres in order to analyze word and style choices. This analysis will be connected to thinking about the intended audiences for certain journalism styles and publications. Finally, students will be asked to reflect on what they learn in the context of their own writing for the course.

Data

The data will be collected, formatted, and provided to the students by the instructor. Correctly formatted data should be text-only and saved as text files. Texts can be copied from the Web and pasted into a basic text editor like Notepad to format. Each text should be saved as its own text file (i.e. if 5 journalism articles are being examined, each should be its own text file).

The instructor will choose what type of journalism writing will be examined for the project (for example, feature articles, hard news, reviews, or columns). Examples of writing will be gathered across publications of various types, sizes, or other variables. The subject matter of the articles should remain constant. For example, 5 feature articles about the same politician appearing in 5 different publications with a range of political leanings could be examined. If the instructor chooses to, s/he may solicit ideas for the content from the students.

Once data is collected and formatted, the files can be distributed to students via Canvas or email.

Students should read and be familiar with each of the texts before completing the analysis. This will allow a more nuanced and educated discussion of the analysis findings.

The entire class could work on the same data set, or the instructor could provide a number of datasets and let the students choose. Students may also work in groups for the analysis and discussion.

Analysis

Before students can analyze the data, they need to become familiar with uploading files to Voyant and with the program’s interface. Sample files outside the corpus of the assignment can be used to do this. The instructor may want to spend class time walking students through the upload process and discussing the different functions and visualizations of the program. The instructor should also discuss the concept of “stopwords” and demonstrate how stopwords affect Voyant’s output.

Once students are familiar with the program, they can upload the dataset to Voyant and begin analysis. The first step of analysis should be consideration of the current list of stopwords. As the default list is generically provided by Voyant, students and/or the instructor may decide that the list needs to be amended. Students should look at the initial output to determine if there are keywords that need to be added or removed from the stopword list. For example, when analyzing features about a particular politician, the name of the politician may not be necessary to include, as it remains constant across texts (however, a case could be made to keep the name in order to analyze its use across texts; decisions to include or subtract words should be reasoned out). Editing the stopwords list can be done from the Options pop-up window.

Once the stopword list is satisfactory, students can begin further analysis. Student should examine:

* The word cloud in order to get a sense of word frequency
* TermsBerry in order to explore colocation of key words
* The summary statistics, including Vocabulary Density, Average Words per Sentence, and Distinctive Words
* The Contexts in which high frequency or Distinctive Words appear

While examining each of these, students should discuss findings with other students (perhaps in small groups, and then in a share-out to the entire class). Ask students to explain what they think that data indicates about the similarities and differences between the texts (i.e. different publications) in the corpus.

Finally, prompt students to make connections between the data and the intended audiences of each text. This will require some previous knowledge of the publications that the texts came from (this assignment shouldn’t be distributed until students are able to do this, which may require a preceding lesson about publication types). Students and the instructor should discuss results of the analysis in the context of how language in the texts is used to connect with audience. The instructor may also have students use an online reading level tool (like [Readability Formulas](http://www.readabilityformulas.com/free-readability-formula-tests.php)) to find additional quantitative info and data about the texts or publications in order to understand the intended or actual audience.

Connection to Own Writing

After completing the analysis and discussing the results with classmates and the instructor, students are asked to reflect on what they learned about style and audience and connect the lesson to their own writing. The reflection may take the form of a short (~2-3 page) essay. In the essay, students are encouraged to respond to these prompts:

* Summarize (in 1-2 paragraphs) the major findings of the analysis. Include details on word choice and style.
* Did anything surprise you about the data or results?
* What were the major word and style choice differences between the texts?
* How consciously do you think about word choice when writing for the media or particular audiences?
* The style of which text most matches your natural or current writing style? Which text least matches your natural or current writing style?
* In what circumstances does it make sense to take risks with word choice and style? In what circumstances does it make sense to write conventionally?