Jeffrey Tharsen tharsen@uchicago.edu

“Advanced Methods for Text Visualization” RCC Workshop

Spring 2021

***Sources:*** [**https://github.com/rcc-uchicago/advanced-text-analysis-2**](https://github.com/rcc-uchicago/advanced-text-analysis-2)

1. **Visualizing Stylometry (HCA Dendogram & *k-means* PCA)**
   1. ***LEXOS* (Comparative Stylometry : Dendrogram + PCA) :** [lexos.wheatoncollege.edu](http://lexos.wheatoncollege.edu/)
   2. ***Python-based Stylometry* : Stylometry\_HCA.ipynb , Stylometry\_PCA.ipynb**
2. **Visualizing Word2vec**
   1. ***Python-based Word2vec* (gensim) : Word2Vec all\_Shakespeare.ipynb**
   2. **Word2vec / Tensorflow Embeddings Projector :** [projector.tensorflow.org](https://projector.tensorflow.org)

*(Uses the files tensorsfp.txt and metadatafp.txt from (a) above)*

1. **Visualizing Intertextuality (Text Reuse / Alignments / “Parallel Passages”)**
   1. **TextPAIR EEBO-ECCO :**   
      [anomander.uchicago.edu/text-pair/eebo\_vs\_ecco](https://anomander.uchicago.edu/text-pair/eebo_vs_ecco/)
   2. **TextPAIR Frantext :** [anomander.uchicago.edu/text-pair/frantext](https://anomander.uchicago.edu/text-pair/frantext/)

**TextPAIR Viewer (TPV) for Frantext vs. the *Encyclopédie* :** [users.rcc.uchicago.edu/~jcarlsen/TPV/TPV\_frantext\_encyc](https://users.rcc.uchicago.edu/~jcarlsen/TPV/TPV_frantext_encyc/)

* 1. **TextPAIR “24 Chinese Histories” :**   
     [anomander.uchicago.edu/text-pair/histories](https://anomander.uchicago.edu/text-pair/histories/)

**TextPAIR Viewer (TPV) for the “24 Chinese Histories” :** [users.rcc.uchicago.edu/~jcarlsen/TPV/TPV\_histories](https://users.rcc.uchicago.edu/~jcarlsen/TPV/TPV_histories/)

* 1. **TextPAIR *Aozora Bunko* :** [anomander.uchicago.edu/text-pair/aozora\_5\_21](https://anomander.uchicago.edu/text-pair/aozora_5_21/)

**TextPAIR Viewer (TPV) for the *Aozora Bunko* :**   
[users.rcc.uchicago.edu/~jcarlsen/TPV/TPV\_aozora](https://users.rcc.uchicago.edu/~jcarlsen/TPV/TPV_aozora/)

“Quote Matching” using BLAST (for literary sequence analysis) :  [pvierth.herokuapp.com/tools/match](pvierth.herokuapp.com/tools/match/) ; [pvierth.com](https://www.pvierth.com/)

1. **Visualizing Patterns in Phonology (IPA)**
   1. **PhoTransEdit :** [photransedit.com/Online/Text2Phonetics.aspx](http://www.photransedit.com/Online/Text2Phonetics.aspx)
   2. **Robert Frost’s “Stopping by Woods on a Snowy Evening” in  
      the Visual Text Explorer (VTE) :** [edoc.uchicago.edu/vte/vte.php](http://edoc.uchicago.edu/vte/vte.php)

*(Use the input files “frost\_stoppingbywoods\_IPA.txt” and “vte\_IPA\_list.txt” ;*

*Select “Vertical Display” and “Split on Characters”.)*