

George Floyd Transcript Data

Dr. Leah C. Windsor
Research Assistant Professor, Institute for Intelligent Systems
Faculty Affiliate, Department of Political Science
The University of Memphis
Leah.Windsor@memphis.edu
<https://github.com/lecauwe/georgefloydtranscript>

Overview

This github file provides the turn-by-turn transcript of the George Floyd arrest [1]. It includes the LIWC 2015 raw data, and summary statistics for all LIWC variables for each of the speakers [2]. This transcript data was derived from the original Minnesota court records.¹ In the data file, the id variable is the speaker turn number, not reflective of timestamp.

Figure 1 shows graphs of the four officers and George Floyd along six variables: Power, Swear, Cognitive Processes, Masculine language, Feminine language, and Perception. Figure 2 shows graphs of the four officers and George Floyd along an additional six variables: Body, Risk, Health, Anger, Words per sentence, and Authenticity.

When referencing this corpus, please cite:

References

1. Windsor L. George Floyd Corpus. 2020. Available: <https://github.com/lecauwe/georgefloydtranscript>
2. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. UT Faculty/Researcher Works. 2015 [cited 9 Dec 2016]. Available: <https://utexas-ir.tdl.org/handle/2152/31333>

¹ <http://www.mncourts.gov/mncourtsgov/media/High-Profile-Cases/27-CR-20-12951-TKL/Exhibit207072020.pdf>

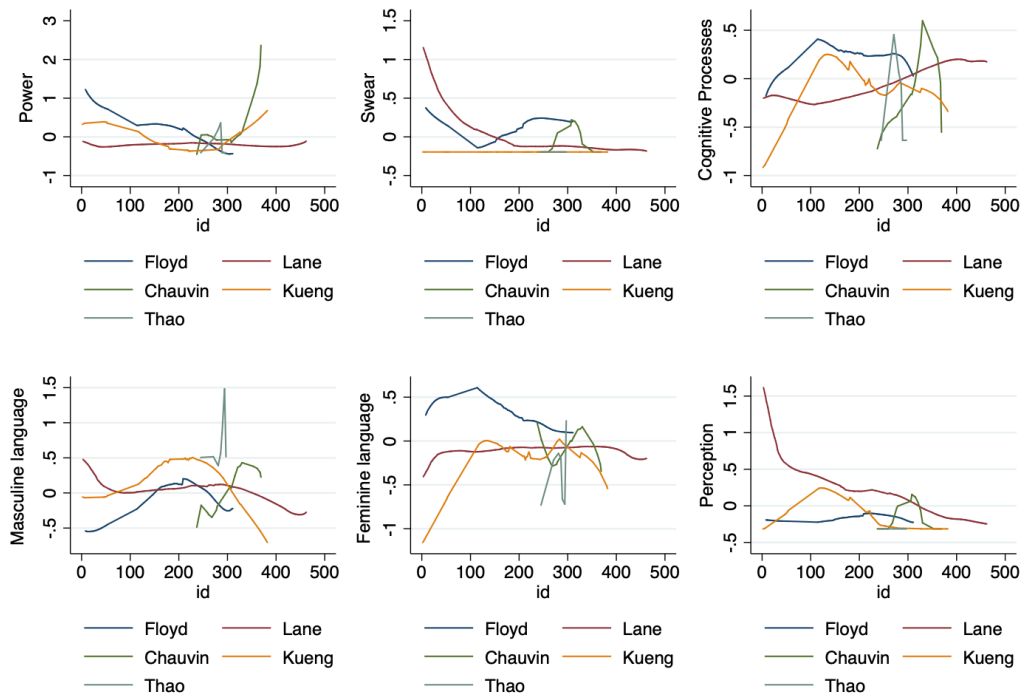


Figure 1. Power, Swear, Cognitive Processes, Masculine and Feminine language, and Perception

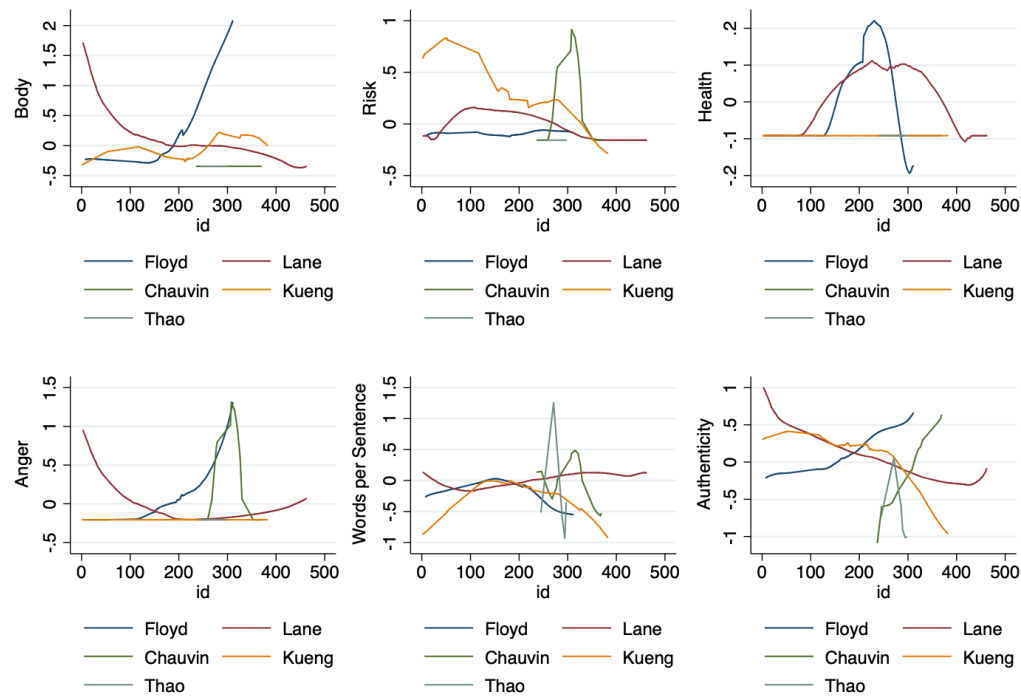


Figure 2. Body, Risk, Health, Anger, Words per sentence, Authenticity