

Reflections on Stylometry, Part I

Reading

- Pre-Class Reflection
 - Whissell, Cynthia. "Traditional and Emotional Stylometric Analysis of the Songs of Beatles Paul McCartney and John Lennon.", *Computers and the Humanities*, vol. 30, no. 3, 1996, pp. 257–65.
 - Holmes, David I., and Judit Kardos. "Who Was the Author? An Introduction to Stylometry.", *Chance*, vol. 16, no. 2, 2003, pp. 5–8.

Clay Glover

1:26 PM Reading Holmes' article Who was the Author was fascinating. I enjoyed learning about how stylometry is used, and how it can be applied to attempt to determine the authors of different works. The article maintained that one of the greatest strengths of the technology is its ability to match the unique structure and style of certain authors to texts, which pre-generative AI was certainly useful. However, with the advent of this technology, which can be fed previous texts and trained on them until it can almost perfectly replicate their author's style, it seems that stylometry will no longer be able to offer this value to researchers. Perhaps there will be ways to solve this issue, yet it appears that the field will be dramatically impacted. I also enjoyed reading Whissell's article that discussed employing stylometric analysis to study Paul McCartney and John Lennon's songs. It was interesting to see how the technology could offer quantitative support for the qualitative observation that Lennon's songs were less upbeat. Using stylometric analysis, critics could trace how Lennon displayed less positive words in his lyrics. While the field of stylometry will certainly evolve in the coming decade, it already offers significant value to researchers, as demonstrated by these two articles.

Raphaella Gold

10:27 PM I really enjoyed looking at stylometric analysis as a digital humanities tool I hadn't considered before. Like other tools we looked at, stylometry could either be digital or non-digital, but digitization would definitely render such precise statistical analysis less complicated. It was interesting to see how stylometry has been criticized as "cold and impoverished, studying words without their meanings" according to the Whissell reading, while also being praised for its "relative objectivity." Both the criticism and the praise felt connected to assessments of other tools we've looked at, like topic modeling, which is also more objective than an individual human, but focuses more on pattern and frequency than meaning. Still, Whissell was able to formulate a new stylometric technique which could incorporate some level of meaning into what had previously been a purely numerical evaluation of words.

I found myself still curious about how in practice, emotional stylometry can bring people closer to making "more complex judgments" about a text. For example, if someone reads and interprets a text on their own, it would seem to me that they would be able to come to the same conclusions and drawn their own inferences about the words. Though I found the argument about stylometry's objectivity convincing, one hole in that argument is that even if the assessment of the words through stylometry is objective, an individual's inferences based on the stylometric analysis would still be subjective.

It was also notable that the researchers chose to remove the musical element of the Beatles songs they studied by limiting their study "to the Lennon-McCartney lyrics which are treated as verbal texts and submitted to emotional and traditional stylometric analysis." Thus, the study almost treats Beatles lyrics as poetry and acknowledges that it does so, which makes applying stylometry to the songs much less challenging. This did make me curious about whether something like stylometry has been created for musical analysis, which tends to be even more subjective than literary analysis. I did a bit of research and found that there actually has been a lot of research into musical stylometry using an approach of pattern recognition rather than word recognition, which I found fascinating.

On a separate note, in terms of the Holmes and Kardos reading, I wrote in one of my annotations that I've noticed that many of our readings have felt the need to clarify that the new technology they promote does not threaten old/traditional methods of scholarship – that close reading will not be eclipsed by distant reading, topic modeling will not eliminate the need for engaging with various forms of media etc.

Holmes and Kardos write, "Throughout this story of stylometry's coming of age, it is important to emphasize that it presents no threat to traditional scholarship." They add that stylometric evidence must be taken "in balance along with that provided by more conventional studies made by literary scholars." This makes me wonder whether in general, defensiveness might be a quality shared among digital humanists. Whether this is because the field has genuinely faced an excess of criticism from more traditional humanities scholars, or due to the anticipation of criticism I am still not sure.

Andrew

2:14 PM I think stylometry is the most fascinating topic I have encountered in this class so far. This is a way in which statistical analysis can find psychological, emotional, and other exciting patterns to do with authorship, answering important questions such as the validity of author contributions. I am deeply fascinated by psychology and how that influences one's style of writing (or is reflected in one's writing), so learning specifically about the uniquely quantifiable qualities of Shakespeare or Lennon and McCartney was thoroughly engaging. I am also a huge fan of the Beatles so I was deeply invested in the differences between Lennon and McCartney according to emotional stylometry. There are ways in which this analysis makes sense and other ways where I had many questions. The depersonalization of their songs over time made sense because it was strictly analyzed through the lower frequency of first and second pronouns – that works in an objective way, but when it comes to emotion and style over time, I feel like there are many other factors. First of all, is there a standardized positive and negative association for words in English? There must be many different dialects across subcultures and slang that vary the emotional context for words. Secondly, how do you account for the cultural shift in America that was very prevalent during the time of the Beatles. Could this have influenced their lyrics to become more depersonalized and negative later on? Lastly, I think that categorizing time as four stages is a little too simplified and I am curious to see how Lennon and McCartney's styles change throughout a more specific representation of time. Overall, I found stylometry the most fascinating type of analysis I have come across recently and it provides new dimensions to the field of humanities that I look forward to hopefully uncover one day.

Anyia Kalogerakos

2:46 PM I really enjoyed learning about stylometry and I think it is one of the digital humanities practices that make the most sense to me. It is very intuitive that any author has a particular style and this style can be measured using word frequency, tone, and other syntactic methods. The Whissell reading was a fun reading, as I am a big Beatles fan and found their methodology and conclusions interesting. I also felt that

the choice of using the Beatles for their study was a clever choice, as it lends credibility to the practice of stylometry since many people are very familiar with the Beatles' songs, members, and overall history. Therefore, the conclusions drawn in this article match with a wide population's (Beatle fans') expectation. The Holmes and Kardos reading was helpful to get some further background on stylometry, but I found it most interesting in its discussion of future methods and uses of stylometry. I am interested in the use of neural networks for stylometry, as it seems to be perfectly suited to this type of difficult-to-see pattern recognition. I am curious how these methods can be applied to "emotional" or "connotative" stylometry that the Whissell reading discussed.

Melissa Woo

2:15 PM I found both of these readings on stylometry quite interesting – it seems like a very intuitive and helpful area of digital humanities. The Beatles and Federalist paper examples were cool case studies to recognize the various situations in which it is interesting to better understand the author's style and how it changes over time. I'm curious about the types of pushback more traditional scholars have levied against stylometry; I know that the Whissell reading mentioned criticism for being an "old and impoverished method of textual analysis", but it seems that the scale at which one could quantitatively analyze text on a word-level could reveal a lot about style in a manner that humans wouldn't want to devote the resources to doing. This seems wholly complementary to traditional analyses (as the end of the Holmes Kardos reading mentioned) rather than being positioned to overtake or replace certain sections of existing research. I wonder how stylometry has evolved since its inception; it seems that technology and machine learning (like neural networks, as mentioned) has the potential to continue improving the method and its results. As I saw mentioned in the Perusall comments, there are also a lot of interesting questions surrounding generative AI and large language models and how these developments might impact stylometry. Can stylometry distinguish AI-generated work (like GPTZero or other Chat-GPT detection style tools) or perhaps also aid in making better contextualized analysis of style? I'm also interested in what scholars say about how collaboration impacts stylometric analysis – if, for example, a collaborator makes word-level edits or revisions, how might this be incorporated into a writer's style and/or change the findings of stylometry applied to that corpus?

Pippa LaMacchia

7:36 PM It was fascinating to learn more about digital stylometry as I love that we are compiling all of these unique ways in which computational analysis furthers our understanding of humanistic fields. The Whissell article was interesting because it illuminated the idea that stylometry transforms objective analysis onto an emotional field of study. In previous articles and discussions it seemed impossible to reckon with analyzing fiction properly through other tools like topic modeling but stylometry provides an answer to this dilemma. I was particularly struck by the application to John Lennon and Paul McCartney's music as a concrete example of how stylometry authentically works. I now understand that stylometry uses concrete literary ideas to communicate more nuanced emotional threads in a way I didn't know was possible. The Holmes' article just elaborated upon these ideas — explaining that stylometry seeks to "complement" traditional scholars and assuage any concerns about technology as a threat to humanistic research.

Alison Fortenberry

9:03 PM I really enjoy the idea of stylometry. A writer's voice is usually something I can pick up on if I've read a lot of their works, but it's hard to qualify why. There are repetitive words and sentence structures, but I think, as a reader, I'm not actively picking up on those elements, and the text just feels familiar. It's

interesting to be able to point to specific, quantifiable elements that help recognize an author's voice. Holmes and Kardos discuss training AI to perform stylometry, and it made me question whether generative AI could complicate stylometry or render it inaccurate. If I asked ChatGPT to write a song as John Lennon, would there be enough stylistically different (assuming ChatGPT produced a good output) to differentiate it from a song Lennon actually wrote? Maybe in its current state, generative AI has enough tells that it's not John Lennon, but assuming it becomes more accurate over time, is this something that stylometry could recognize? Is it possible for stylometry to adapt to expanding text authorships with the inclusion of AI, or does that go beyond what it could detect? Will generative AI render stylometric methods useless (or heavily devalued if a writer's voice can be easily mimicked by AI, thus complicating its ability to determine authorship), or will it introduce a new wave of stylometric interpretation?

James Sowerby

11:19 PM The question of authorship has come up a lot so far in a few of my other classes. When reading Aristotle's *Poetics* and *Politics* together, we discussed how scholars spend inordinate amounts of time combing through the lines to find spurious additions from later commentators, often using other surviving manuscripts to judge between versions. In these examples, it's hard to build a sense of style to train models along the lines of the Holmes article since there are so many variations. In fact, I'm not even sure that authorship can totally be established in these antique cases, since there have been translations upon translation that have preserved them over time. The *Poetics*, for example, were translated into Syriac, then Arabic, then Latin—the Greek original does not exist. Ignoring the fact that these stylometric tools seem to be focused on English, the statistically likely Aristotelean texts in one edition are probably very different from the first version. That aside, it's a super interesting question and I find myself very taken with the potential tool of stylometry. I was surprised to see that these articles were written so long ago. Laure Thompson mentioned that the jsLDA method was very old—which is unusual in computer science—but these seem just as old as both articles are over twenty years old. As usual, I remain dubious that stylometry is as reliable as humanistic methods. There seem to be too many potential issues that would present roadblocks—to what degree are the quirks of writers figments of their time period, and not of their individualistic style. I'm also thinking about how stylometric tools might themselves create style recursively. I think I maybe mentioned Adorno's *Culture Industry* before, but one of his theses for resisting the totalizing fascism of prescriptive style and media are independent works of art, namely those that resist stylometric characterization and cannot easily be classified. Maybe it's a reach, but I couldn't shake it as I read both of the articles for today.

Emanuelle Sippy

11:36 AM Whissell's piece really helped to showcase what emotional stylometric analysis added to traditional stylometry. Distinguishing features of McCartney and Lennon songs, as well as tracking "developments" in their music at various stages illuminates the value of these more objective methods in relation to traditional literary studies. However, I also question how subjective the Dictionary of Affect ratings are. It seems like determinations about how "active" or "pleasant" a word is would vary contextually and based on the author's own subjectivity. I would like to better understand how these ratings are established. Additionally, I thought there was an interesting tension in the Holmes and Kardos article about whether "rare words" or "common function words" are more likely to indicate an author's literary style. I am curious about how experts in the field have weighed this and/or if models have been built that seek to account for both rare and commonplace words. Lastly, Holmes and Kardos explain that AI techniques in stylometry necessitate "sufficient data" of other identified works from an author's corpus (7). While this makes perfect sense, it seems like this requirement would pretty severely limit the authors that AI

techniques in stylometry can evaluate (i.e. more well-known authors who tended to produce more published works). I wonder about how this limitation might affect the representation of the authors that AI techniques in stylometry can analyze (from period to geographic location to identities) and how researchers in the field might address this limitation.

Pia Bhatia

3:27 PM The piece by Holmes and Kardos made me think of so many different works of art I have interacted with. I had so many questions in response to popularizing stylometry, including: what does it tell us about art forgery, and is it more acute than specialists in certain regards? And is stylometry able to detect universals in an artist's 'style' across a lifetime of their work developing? I would love to hear what people think about this in our next discussion. I also think this tool could be especially useful in noting certain trend preferences over time — a piece I read (<https://www.thedrifthmag.com/controlled/>) about 'the millennial sex novel' as a kind of genre that has developed today caused me to think of what we deem to be the parts of a text that make a genre or a movement. Perhaps, beyond contemporary features in the content of certain works, specific style choices are favored today as well — I have often heard people lament how many formal elements of poetry are lost in the poems that are popular today, and perhaps stylometry could elucidate the extent to which this is true. New

Talia Goldman

10:04 PM I enjoyed considering ideas of authorship when reading Holmes and Kardos' "Who Was the Author" and Whissell's "Traditional and Emotional Stylometric Analysis of the Songs of Beatles Paul McCartney and John Lennon." Discussions of AI "impersonating" certain authors immediately came to mind, and thinking about stylometry as a way to ascertain the emotional patterns within an author's writing, which is closely tied to their personal experience and social/cultural context, further problematized AI "writing" in styles almost exactly similar to real authors. I wonder if stylometry will fuel AI's ability to replicate writing styles, or if it will be used to further confirm/analyze the more "authentic" human emotion of authors. Like others have noted, I was also very compelled by a DH tool that actively incorporates emotion, which seems particularly consistent with a defining feature of the humanities—to deal with large questions, sometimes ambiguous, that reveal aspects of the human condition through time. Learning about stylometry shifted my thinking about quantitative tools being a bit too "impersonal" sometimes, contrary to Whissell's comment that stylometry is sometimes seen as "cold" (Whissell 257). I think stylometry would be quite compelling to humanists, but, of course, it does have potential to be misused (especially in regards to AI), and data coming from stylometric analysis should be used in conjunction with historical evidence on the analyzed text and its author, as demonstrated in the Beatles example.

Helen Gao

11:31 PM From the Beatles stylometric analysis, it seems that stylometry seems to parallel a lot of the other techniques we've read about in the sense that it's good in combination with subjective scholarly observations as an objective way to support existing claims, but not sufficient for analysis on its own. Though stylometry was described as being more objective, it seems that measures like the Dictionary of Affect still introduce some level of subjectivity, as participants in many studies tend to be "WEIRD" and thus their interpretations of certain words may not reflect how the broader population would read the same passage. The proposed emotion clock was also interesting, and provided a nice visual representation of the style of the songs. Additionally, I was curious about how they chose the 'hours' for the clock (both in terms of which words they chose and how many words they chose). "Who Was the Author?" also reiterated that

stylometry “presents no threat to traditional scholarship”, once again showing how people writing about the digital humanities tend to include sections assuaging the apparent fears of other scholars. The section about artificial intelligence made me a little concerned about explainability (as every discussion about AI seems to do). Even if we don’t treat the results of neural networks as sufficient evidence for authorship, I’m worried that publishing the results of a neural network analysis of a work with contested authorship would lead to later research suffering from some sort of confirmation bias. Finally, I thought the story about stylometry being applied to legal cases was interesting, because it went beyond the more traditional literary applications that were discussed in the readings.

Colin Brown

12:04 AM From this set of readings, it seems like the gold of stylometry is how it can be applied, interpreted, and utilized. This especially stood out to me in the Holmes and Kardos reading, where I was quickly reminded of current efforts to make detection tools for AI-generated texts. Many think that text from tools like ChatGPT have a distinct writing style that can distinguish them from human writing, so there have been some tools that do a similar task to stylometry. Interestingly, the reading talks about neural nets as a method for better detecting trends in writing; now, two decades later, neural nets are the instruments making the writing that we’re analyzing. I found Whissell’s application of stylometry quite fun. It was neat to see how stylometry can span everything from Shakespeare to the Beatles. Additionally, it was notable to me that the study split the lyrics between the two writers first and then performed the analysis. It could perhaps be interesting to validate the stylometry results by going the opposite direction: take the stylometry results, remove the writer label, and perform unsupervised machine learning to see if it correctly groups the lyrics back into McCartney and Lennon categories. It could be a test of both algorithmic methods.

Layla Williams

12:25 AM Stylometry is an interesting tool system in the digital humanities because it tries to balance the emotional analysis with the quantitative elements of humanities research. I think the example of the Beatles song analysis was especially helpful for understanding where stylometry comes in in the research process. You come in with an established base of knowledge for what thesis you might be trying to test or prove, and you then use the stylometry to back that thesis up. Like a lot of the tools we have used (voyant tools and topic modeling), you are not necessarily using the tool in order to discover questions. It can be a valuable source for proving your points of analysis. With the example from the reading it made it seem like the thesis was based on a larger body of knowledge outside of the scholarly sphere (this was something that might have also been popularized in popular culture). This means that potentially stylometry could be used to prove theses that might initially go against the traditional understanding of the corpus that you might be studying. Rather than confirming a suspicion, you can propose a suspicion to explore and then confirm your suspicions. I do not know if I would personally use stylometry, but I can see its potential uses.

Ethan Haque

5:01 AM The Whissell paper is pretty cool. It's about 30 years old, so some of the language they use and terminology is a little weird. As far as I could tell they didn't really provide enough detail to reproduce the analysis they were doing and I'm still not really sure how they went from the song to the broad conclusions about Paul McCartney and John Lennon. The tokenization process and the details of the text vectorization process are muddy. The emotion clock is a cool visualization to quantize a song and put it on vector subspace. Not sure why they've normalized it to have length 3, though. All in all, super cool to be doing this kind of analysis three decades ago. I felt the Holmes and Kardos article was much more fun to read. I don't

think they are right about the holy grail of stylometry and that it's probably a bad idea to strive for something like that. Stylometry is probabilistic and not deterministic. So, the best we can do is say with a certain amount of confidence that a particular piece was authored by author a as opposed to author b. Take for example the new AI content detection tools like GPTZero. They are wrong so often that it is astounding how they market themselves. In these plagiarism detection contexts false positives are extremely damaging, but false negatives make the tool generally unreliable. Part of this is because these tools require "sufficient data" as Holmes and Kardos put it. There is simply no way to confidently discriminate authors from small sections of text. Even with large amounts of text, there are bound to be people with similar styles just be chance.