Exercise 3: Stylometry

Timing. Students can begin the exercise in class, complete it outside of class, and report their conclusions in a subsequent class.

Tools required. List Words, Excel, and a digital archive such as Project Gutenberg.

Purpose. In the third exercise, based on the techniques described in Jockers' chapter on stylometry, students will use the basic techniques of "literary fingerprinting," including the frequency of signal words, to examine the correlation between the public persona of an author and the style manifested in her work. Does an author's style vary when she writes anonymously, publishes under her own name, or adopts a pseudonym? Is style, as Jockers suggests, deterministic, like fingerprints, or is it adaptable to an author's purpose, like the skin color of a chameleon? In what sense can style be measured quantitatively? Primary case studies could include female novelists of the eighteenth and nineteenth centuries who published anonymously or pseudonymously, including Jane Austen, Charlotte Brontë, Emily Brontë, and Louisa May Alcott.

Step 1. Practice reading literary fingerprints by asking students, individually or in teams, to examine the fingerprints of ten canonical American novels from the nineteenth century (included as a separate document, "Stylometrics of American Novels 2," in the module resources). Referring to Jockers' discussion of stylometry, explain the meaning of each component of the fingerprint, including the density score and the comparative value of each signal word. Call specific attention to the gendered pronouns in each fingerprint, recalling Jockers' argument about determined gendered styles.

Step 2. Tell the students that the unidentified novella in the last column was written by the author of one of the other fingerprinted novels. Ask them to try to identify the author of this novella by comparing its signal word values to those in one of the other fingerprints. Encourage them to use a process of elimination in order to narrow the possible candidates. Given the signal word values, is the novella more likely to have been written by a man or a woman? Is it more likely to have been written in an earlier or later decade?

Step 3. Discuss the students' conclusions. The unidentified fingerprint belongs to *Behind a Mask*, written by Louisa May Alcott under her pseudonym A.M. Barnard. In my

own experience with this exercise, the majority of students have been able to identify it correctly by comparing the signal word values. Ask the students to articulate their reasoning process. Even wrong answers can be instructive, if students can show that they have made sound comparisons and deductions. The similarity of Alcott's style to that of Barnard suggests, as Jockers implies, that an author cannot adapt style to different personae or purposes.

Step 4. Now that the students have practiced reading literary fingerprints, they can practice making them using the process described in detail in "How to Fingerprint a Novel," a separate handout included in the module resources. This process is basically the same one that Jockers uses in his chapter on stylometry. It can seem tedious to students, depending on the number of glitches they encounter and their level of familiarity with Excel functions, but it will be crucial to completing the exercise. Encourage them to work through the glitches, explaining that the process becomes easier when they learn to complete it successfully.

Step 5. Discuss with students, either as a class or in small groups, how they might use literary fingerprinting to answer questions about the relation between authorial identity and style. Help them to form original research questions that they can work to answer outside of class. While most students will follow Jockers and focus on questions of gender and style, encourage them to use literary fingerprinting to interrogate style as a function of other categories such as period and genre. Ask them to report on their findings in the following class.

Conclusion. Many students will conclude, with Jockers, that style is mostly determined by fixed categories including the gender of the author, the historical period in which the text was produced, and the genre of the text. More interesting results, perhaps, will focus on literary "chameleons," writers who show the ability to adopt contrasting styles according to their persona or purpose for writing. In a larger critical context, the exercise invites students to reconsider the relative determinism or adaptability of literary style.