Looking for the Slave's voice: Dictated Slave Narratives and their Amanuenses

Marie-Pierre Baduel

Looking for the Slave's voice: Dictated Slave Narratives and their Amanuenses.

Marie-Pierre Baduel

Université de Toulouse Jean Jaurès – marie-pierre.baduel@univ-tlse2.fr

Abstract

There is, in the English-speaking Atlantic world, a unique literary genre unknown to any other culture: the narratives of ex-slaves. However, defining slave narratives is not easy and researchers do not necessarily agree on the corpus. On the one hand, there are texts written by the ex-slaves themselves, very few of them because not many slaves knew how to write. On the other hand, there are dictated slave narratives, whose place in this particular genre is more problematic. Since they did not know how to read and write, they asked an amanuensis to do it for them and this transcription was more or less faithful to the dictation of the slave. This is the reason why they are sometimes considered as biographies. I decided to include them in my corpus of slave narratives but by doing so, I did not put aside the thorny issue of the amanuensis's influence on the text. Amanuenses cannot have been mere recording machines but I want to see if the slave's voice can nonetheless be heard. We will never know for certain to what extent what we are reading is a reflection of what the ex-slaves said but new tools such as textometry and authorship attribution can help us demonstrate that the slave's voice is present. I am presenting here a case study of three slave narratives which are said to have been transcribed by the same man. To this end, I used the IRaMuTeQ software (created by Pierre Ratinaud, LERASS) and all its functionalities to find differences between the narratives and the text the amanuensis wrote under his own name.

Keywords : slave – amanuensis – narrative – dictated – IRaMuTeQ – authorship attribution

Introduction

For my PhD in American history, I study slave narratives written before the American Civil War. But what is a slave narrative exactly? The answer to this question has been controversial for decades now and it has not been resolved yet, and probably never will. There are, on the one hand, narratives that were written by the slaves themselves, about 25 of them have been proven authentic and thus are not questioned by researchers, although we can still wonder how much influence white abolitionists exerted on their writing. On the other hand, there are dictated narratives and here the debate begins. A large majority of slaves did not know how to read or write because teaching them was forbidden in the Southern states. As a consequence, those who wanted to tell their story asked someone else, usually white men, respected in the community but not necessarily abolitionists, to write down what they were telling them. Of course, the amanuenses insisted in the prefaces or introductions that the narrative was "written under his dictation" (Brown, 1855) or "taken mainly from her own lips" (Elizabeth, 1863) to insist on their authenticity. There are, for the moment, 24 such narratives in my bibliography.

We will never know how faithful the amanuenses were to the words they heard but it is a crucial question because it is what draws the line between a dictated narrative and a biography. Therefore, we need to find ways to "hear" the slave's voice in these narratives. Thanks to Pascal

Marchand and Pierre Ratinaud[1], I was introduced to the world of textometry and I thought that one aspect in particular, authorship attribution, could be helpful in this quest. I used the software IRaMuTeQ (created by Pierre Ratinaud, from the LERASS).

1. A Case Study

2.1 Characteristics of the Corpus

To try and determine who the real authors of the dictated slave narratives are, we need another text or other texts to compare them to. Luckily, most amanuenses also wrote books under their own name.

For the present case study, I chose three narratives: Lewis Clarke (24,692 words), his brother Milton Clarke (9,929 words) and James Matthews (14,671). The Clarke brothers' narratives were transcribed by the same person, Joseph C. Lovejoy and, according to Susanna Ashton, he is also "quite likely" Matthews's amanuensis (Ashton, 2014). I also found a speech delivered by Lewis Clarke and written down by Lydia Maria Child (6,551 words). Lovejoy wrote a preface for both Lewis's and Milton Clarke's narratives and I also found two speeches he delivered about prohibition (total: 13,359 words). I separated these from Matthews's preface since we are not completely sure it is the same person. This preface is 1,240 words long and therefore too short to bring significant results to the present study.

There are three advantages in this choice: I can compare three slave narratives which contain the same lexical fields and the fact that the same person transcribed two ("quite likely" three) texts can show us if we hear different "voices" or if the narratives are homogeneous. Moreover, Lewis and Milton Clarke were brothers; consequently, although they were separated when they were children and reunited once free, they shared a part of their history.

First, I shall consider where they are compared to the other dictated slave narratives: they are not really apart from the others, but they are not close either. I calculated the correspondence analysis of the texts (comparing the vocabulary of the different texts, figure 1) and Labbé's index (figure 2)[2]. "This index allows to assess to what extent two texts are close to or distant from the point of view of the lexicon that compose them." (Ratinaud, 2018). Two of the dictated slave narratives stand out from the rest of the texts (Isaac Jefferson because he is proud to be Thomas Jefferson's slave and Peter Wheeler because he devotes only one third of his text to his life as a slave). Consequently, for the correspondence analysis, I did not consider them to make this comparison and thus see the differences among the others better. Here IRaMuTeQ takes all the forms present in the corpus.

[1] I would like to take this opportunity to say how grateful I am for their help and advice.
[2] We can notice that the graph was the same whether I chose lemmatization or not.

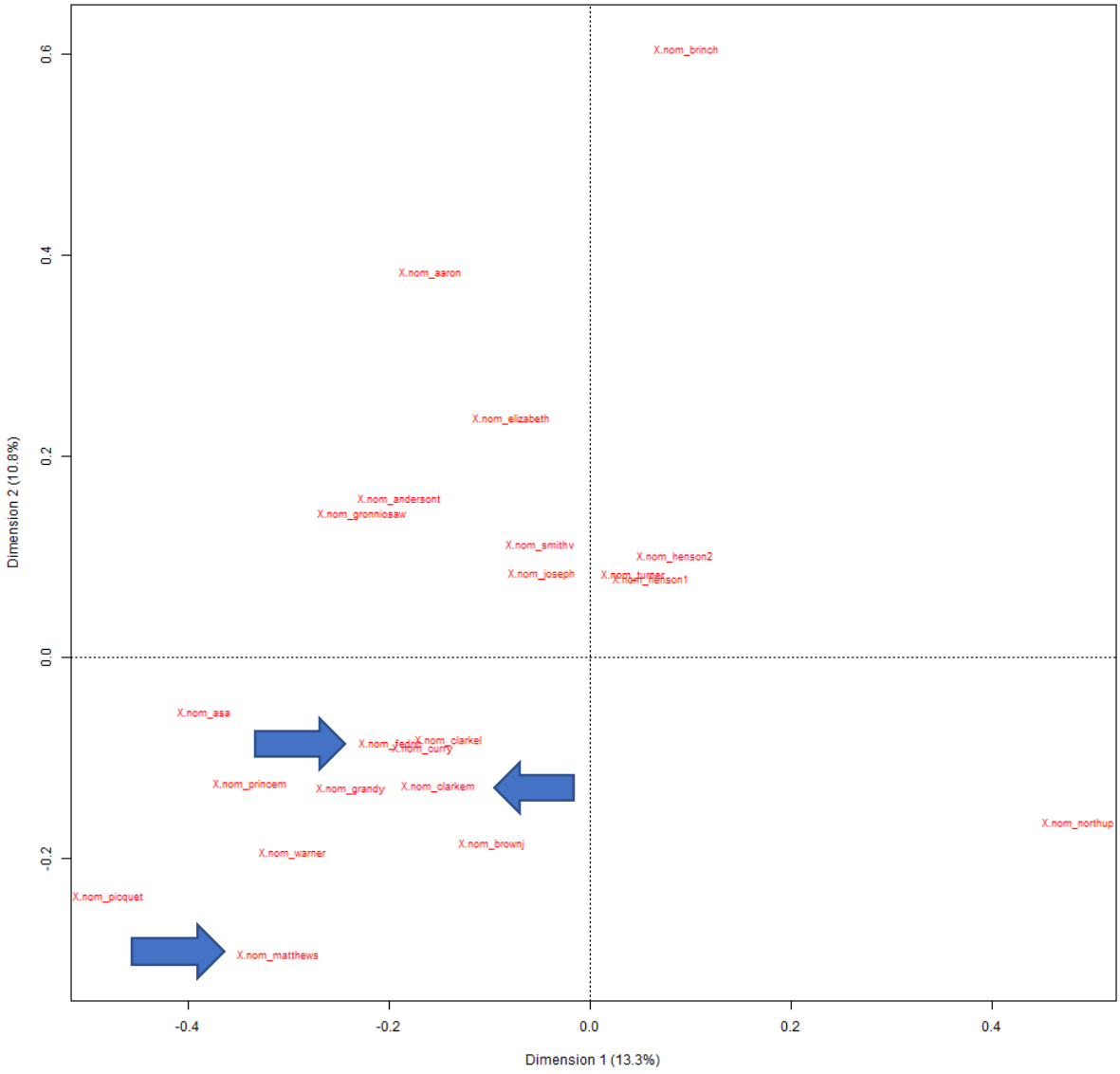

Figure 1: Correspondence analysis of the 22 dictated slave narratives (Wheeler and Jefferson excluded). The arrows point to the three narratives in my case study.

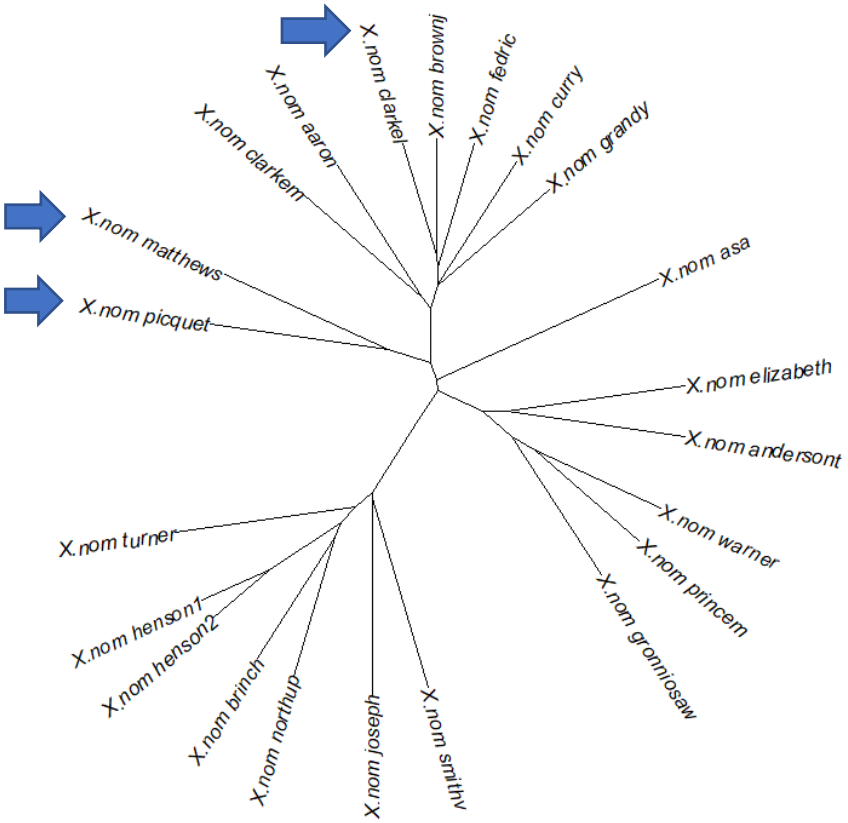

Figure 2: Labbé's index of the 22 slave narratives (Wheeler and Jefferson excluded).

2.2 General Comments

Using the Reinert classification (figure 3), IRaMuTeQ finds 7 lexical fields in my corpus: the most important one is group n°7 (16.6%). It is mainly composed of legal terms and deals with alcohol prohibition, which is the topic of Lovejoy's speeches. Then come group 1 (18%) about the harsh treatments the slaves endure, and group 2 (15.6%), which is more about the environment and the slaves' everyday lives. Group 4 (14.3%) is about discourse and speech. The word "money" is in this group because it is often linked with indirect speech. Group 5 (13.2%) clearly deals with family while group 6 (15.5%) is more difficult to interpret but on the whole it is about running away and life after slavery: Lewis Clarke ran away by following the Ohio River. Milton Clarke was allowed to go up and down the river with a written pass and he used this opportunity to escape as well. The friends are evidently the people who helped them escape, etc. Finally group 3 (10.6%) is about slavery as an institution. We can notice that the 7 lexical fields are quite similar in size. However, the biggest group is the one dealing with alcohol prohibition although Lovejoy is the only one raising this issue. It may mean that he is also the one using the most varied vocabulary to talk about this subject. I will come back to this point later.

Figure 3: Reinert classification of the six texts in our corpus.

When we look at where the different texts are distributed in these groups (figure 4), we notice that Lovejoy is clearly apart from the others because he talks about prohibition and religion (he was a minister). What is interesting is the difference between the slave narratives because the narrators apparently all talk about the same subject: their lives as slaves. James Matthews talks extensively about his work as a slave but also about being whipped very often (group n°2). Lewis Clarke chose to talk about slavery as an institution in his speech (group 3) and not about his life as a slave, which was the main topic of his autobiography. If the three ex-slaves do not focus on the same themes, it may mean that they were rather at liberty to choose what they wanted to talk about. We know that some amanuenses or editors (people who helped ex-slave writers edit their text for publication) did modify the language used and influence the content. For example, Lydia Maria Child, the woman who transcribed Lewis Clarke's speech in my corpus, also prepared Harriet Jacobs' *Incidents in the Life of a Slave Girl* for publication and we can read in her letters to Jacobs what she asked her to do: "My object in writing at this time is to ask you to write what you can recollect of the outrages committed on the colored people in Nat Turner's[3] time … Please write down some of the most striking particulars, and let me have them to insert" and "I think the last chapter, about John Brown[4], had better be omitted" (Yellin, 2008, p.279).

[3] Nat Turner was the leader of one the most important slave rebellions in the South in 1831. His narrative is also in my corpus of dictated slave narratives.

[4] John Brown was a white abolitionist who believed in armed rebellion to end slavery and free slaves. He is best-known for his raid on Harper's Ferry in October 1859. https://www.history.com/topics/abolitionist-movement/john-brown (consulted on 21/12/2019).

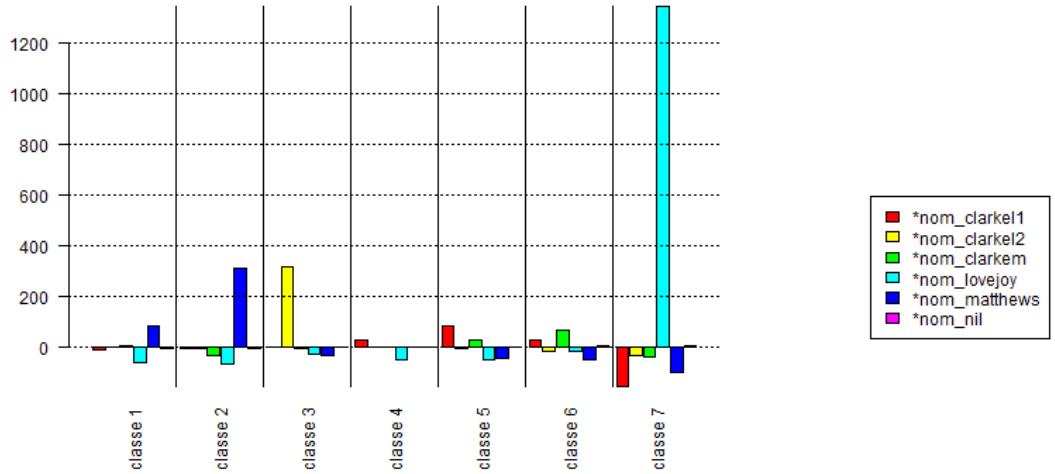

Figure 4: Distribution of the Reinert classification according to a variable (here the name of the narrator).

On the lexical level, I looked for differences between the narratives because I think that these differences, if any, show that the slave's voice is present, not only in the facts told, but also in the vocabulary used to tell those facts. There is not one criterion which can make us conclude for certain that we hear this voice in the narratives. Nevertheless, several criteria might give us an idea.

In the 725 words contained in our corpus[5], there are 49 three-syllable words and 18 of them are mainly present in Lovejoy's text (15 in Lewis Clarke's narrative but the text is almost twice as long). What is more, if we look at words that are only present in one of the texts, we find 20 words, 16 of which are in Lovejoy's text, most of which are two- or three-syllable formal words. For some of them it is logical given the theme of his speeches (for example "abstinence", "fermented", "prohibitory") but some of them are less obvious ("enforce", "blessing", "destruction"). Patrick Juala notices that "word lengths (De Morgan, 1951) might actually be informative. They have long been considered to be a mark of intelligence and / or education…" (Juala, 2018). The comparison between a highly literate man and three illiterate ex-slaves tends to confirm this idea. Labbé's index also confirms a strong difference between Lovejoy and the ex-slaves[6].

If we look at the vocabulary used more in detail and more specifically through one feature of IRaMuTeQ, "correspondence analysis," we can notice that there are 53 words which are over-represented in Lovejoy's speeches: there are of course all the words related to alcohol and prohibition since it is the topic of his text ("law", "wine", "prohibitory", "liquor", "rum"…) but what is striking is that 10 words out of these 53 have three syllables, a sign of complexity. Some words could have easily been used by illiterate ex-slaves ("state", "sir", "total", "god", "use"…). What is more, in 11 cases the words over-represented in Lovejoy's corpus are under-represented in Matthews's narrative. The words that are over-represented in the latter are simple, one-syllable words ("whip", "get", "cotton", "wood", "house"…). Matthews clearly

[5] In correspondence analysis, IRaMuTeQ only lists words that appear more than 10 times in total in the whole corpus. Names, numbers and quantifiers were excluded.

[6] For lack of space, we cannot insert the graph here.

stands out compared to Lovejoy or even the Clarke brothers. The three ex-slaves talk about slavery, but the vocabulary used is different.

The number of hapax legomena (the words used only once in a text) is often considered as a marker of lexical richness[7]. Jacques Savoy talks about hapax and dislegomena as a way to attribute authorship although he thinks these measures, supposed to be constant in a given author and different from one writer to the next, can be unstable (Savoy, 2012). Here Milton Clarke's narrative stands out with 53.09% but Lovejoy's text is not very far with 52.10%. Lewis Clarke's speech comes third (45.73%). Lewis Clarke's narrative is close with 44.32%. Matthews is once again different, with 39.06%. Looking at dislegomena (the words that appear only twice) also brings interesting results: Lewis Clarke's, Lovejoy's and Matthews's texts are above 10% (13.52%, 12.2% and 12.07% respectively), Lewis Clarke's speech and Milton's narrative being close to 2%. The great variation between the figures may indicate a different narrator for each text. An indicator of richness is also the number of different forms divided by the number of occurrences in a text. Since this comparison is telling only if the corpora have the same number of words[8], we can only compare Lovejoy's text (13448 words) and Matthews' (14661) and the difference is clear: 18.79% for the former and 11.80% for the latter.

2.3 Comparison: Function Words

According to Zhao and Zobel, "the appeal of function words is that they are a marker of writing style" (Zhao and Zobel, 2005). We think that this is a key feature in the present study, because the use of function words is mainly unconscious, the authors use them without thinking, contrary to the lexical words which can be carefully chosen. We know that amanuenses or editors transformed what they heard into written language, and therefore some changes had to take place and sometimes it is more obvious than at other times: for example, when Thomas Pringle, editor of Asa Asa's narrative, quotes the ex-slave, he uses vernacular language ("Me no father, no mother now") but in the narrative itself, transcribed by Susanna Moodie, the vernacular language disappears (Asa Asa, 1987). Lovejoy explains, in his preface to Lewis Clarke's narrative, that "much of it is in his own language, and all of it according to his own dictation" (Clarke, Lewis, 1845). In the preface to Milton Clarke's narrative, he does not make any comment on the way he transcribed the narrative. As for the preface to Matthews's text, the editor writes "I have fore, there [sic], as nearly as possible, given his own words". We cannot take what he says at face value, but the differences between the three narratives tend to show that it is true.

Cyril and Dominique Labbé questioned the hypothesis that function words were useful in authorship attribution: "the results of the calculation suggest that, at least for most [function] words, the contrast between authors is reduced" ("les résultats du calcul suggèrent que, au moins pour la plupart des vocables, les contrastes entre les auteurs sont assez faibles"). They disagreed with "anglo-saxon" statisticians who promoted the idea that authorship attribution should concentrate on these words (Labbé and Labbé, 2004). However, function words should be part of the present study because, as stated before, I compare the language of illiterate ex-slaves who never went to school to the words of a highly literate man. Function words can be the sign of complex sentences, sentences that illiterate slaves were unlikely to use. One method

[7] IRaMuTeQ training by Pierre Ratinaud and Pascal Marchand, Jean Jaurès University, Toulouse, 20/11/2018.
[8] Id.

is not sufficient to claim that a text is an author's rather than another's, but we can notice that several methods point in the same direction.

To compare function words (figure 4), Zhao and Zobel's list ("for, in, is, of, that, the") was used ("a" and "and" did not bring significant results) and to it was added all the words that were over-represented in Lovejoy's text. The only function words over-represented in the slave narratives are "I" and "my" for Lewis Clarke, "they" and "he" for Matthews, none for Milton Clarke and finally "she", "he", "if" and "you" for Lewis's speech. Personal pronouns are highly related to the content, so they are not devoid of meaning. Lovejoy does not use many personal pronouns because he does not tell a story, he explains why he is in favor of the repeal of a law prohibiting alcohol[9]. The fact that function words are mainly over-represented in Lovejoy's text shows how rich and complex his discourse is compared to the narratives, and I think that he could not have completely erased that complexity when transcribing the slaves' words.

	Lewis's narrative	Lewis's speech	Milton	Lovejoy	Matthews
of	1,3196	-6,5618	0,4179	33,8974	-39,015
the	-6,0624	-9,7921	2,0714	33,3524	-6,61
its	-6,2113	-1,256	-3,6075	25,5248	-5,3067
this	-0,3162	-2,4772	-0,5484	23,5757	-21,1103
that	-2,4008	1,2575	-2,4433	19,7389	-7,0684
it	-13,9448	1,3973	-7,4591	17,8968	0,9654
which	-0,6828	-3,7585	-1,5986	13,4348	-6,0557
be	0,3665	-2,7987	-0,8104	12,0994	-5,8302
ought	-2,4968	0,2907	-1,325	10,3634	-2,4443
will	-0,6443	0,5918	-1,6538	9,3917	-7,0811
in	0,541	-2,7576	-1,593	7,1664	-2,0917
by	0,3542	-6,7722	-1,1907	5,532	-0,8519
yet	-0,843	0,469	-0,9605	5,0806	-2,7321
your	-1,3643	0,5997	-0,9204	5,6547	-1,9408
under	-2,0053	-2,2148	-0,2515	5,5956	-0,6117

Figure 5: Distribution of some function words in the different texts (more than +5 or less than -5 are considered significant results).

As suggested by Arjuna Tuzzi and Michele A. Cortelazzo (Tuzzi and Cortelazzo, 2018), I also calculated Labbé's index with only function words (figure 6): Lovejoy clearly stands out and Matthews is half-way between Lewis's and Milton's narratives on one side and Lewis's speech on the other.

[9] This may seem surprising but his main argument is that there are numerous references to alcohol and wine in particular in the Bible and therefore alcohol is sanctioned by God and should not be prohibited.

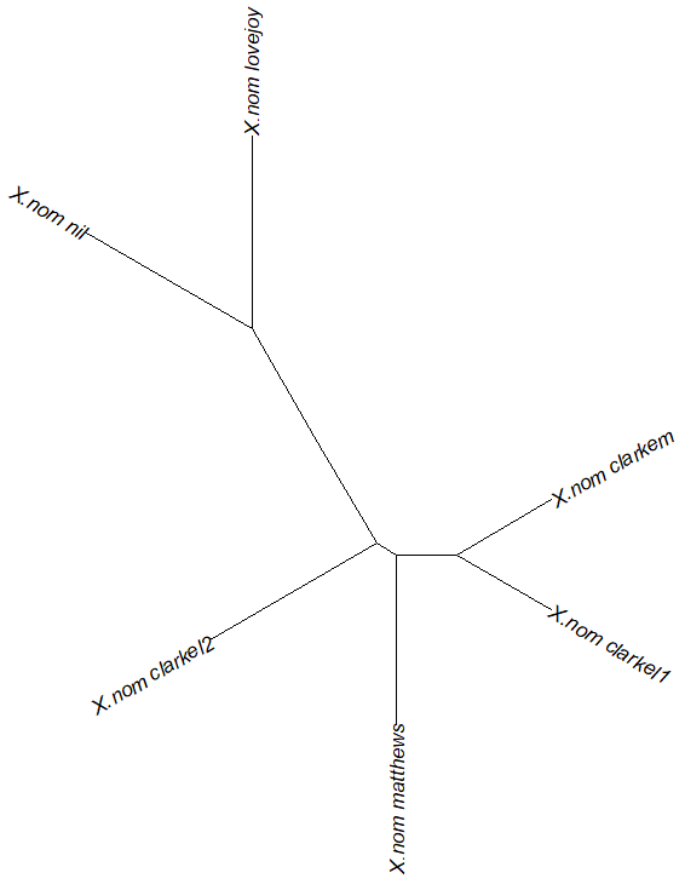

Figure 6: Labbé's index with function words only.

Conclusion

Many calculations point to a real difference between the amanuensis's text and the ex-slaves', on the one hand, and among the ex-slaves' narratives as well, with James Matthews clearly apart from the other two. These differences tend to confirm the hypothesis that the ex-slaves' voices are not erased from the written version of their autobiography. I think that it also shows that textometry can be a very useful tool in my study of slave narratives, dictated and written, so I plan on using those tools for 17 other dictated slave narratives, because the name of the amanuensis is known and other writings from these amanuenses were digitized. In two other instances, the ex-slave (Josiah Henson and Charles Ball) dictated several versions of their autobiography, and thus a comparison of these versions is possible to see how the different amanuenses influenced the text. It would also be interesting to isolate the words of the slave and the words of the transcriber inside the same text when it is obvious, on reading them, that the two co-exist next to each other (Nat Turner's narrative for example) but that seems very difficult. Finally, there are two narratives written by the slaves themselves (Harriet Jacobs and her brother John) but influenced by the editor for which we want to compare letters from the slaves and the editor to the texts. Textometry thus opens many perspectives on a research and a debate that have divided researchers for decades, bringing promising tools to a field where so much can be done still even though some questions will never be answered definitively.

References

Asa Asa (1987). *Narrative of Louis Asa-Asa, a captured African* in Gates, Henry Louis Jr., éd. *The Classic Slave Narratives*. New York: Penguin Books. p.239.

Ashton, Susanna. (2014). "Tales from the Vault". https://www.common-place-archives.org/vol-15/no-01/tales/#.XeuuY_xCc2y (consulté le 19/07/2019).

Brown, John, édité par Louis Alexis Chamerovzow. (1855). *Slave Life in Georgia: A Narrative of the Life, Sufferings, and Escape of John Brown, a Fugitive Slave, Now in England*. London: [W. M. Watts]. https://www.docsouth.unc.edu/neh/jbrown/jbrown.html (téléchargé le 13/07/2019).

Clarke, Lewis. (1842). "Leaves from a Slave's Journal of Life". https://www.docsouth.unc.edu/neh/clarke/support1.html (consulté le 7/12/2019).

Clarke, Lewis Garrard. (1845). *Narrative of the Sufferings of Lewis Clarke, During a Captivity of More than Twenty-Five Years, Among the Algerines of Kentucky, One of the So Called Christian States of North America*. Boston: David H. Ela, Printer. https://www.docsouth.unc.edu/neh/clarke/clarke.html (téléchargé le 8/07/2019). viii.

(Clarke, Lewis Garrard, and) Clarke, Milton. (2001) *Narratives of the Sufferings of Lewis and Milton Clarke, Sons of a Soldier of the Revolution, During a Captivity of More than Twenty Years Among the Slaveholders of Kentucky, One of the So-Called Christian States of North America* in Bland, Sterling Lecater Jr., éd. *African American Slave Narratives: an anthology*, Westport: Greenwood Press. pp.123-191.

Elizabeth. (1863). *Memoir of Old Elizabeth, a Coloured Woman*. Philadelphia: Collins. https://docsouth.unc.edu/neh/eliza1/eliza1.html (téléchargé le 23/07/2019).

Juala, Patrick. (2018). « Thesaurus-based semantic similarity judgments: a new approach to authorial similarity?" in Tuzzi, Arjuna and Michele A. Cortelazzo. *Drawing Elena Ferrante's Profile: Workshop Proceedings, 7 September 2017*. Padova: Padova University Press. p.51.

Labbé, Cyril et Dominique Labbé. (2004). « La Distance Intertextuelle ». *Corpus, 2*. 2004. https://journals.openedition.org/corpus/31 (consulté le 6/12/2019).

Lovejoy, Joseph C. (1853). "Speech of Rev. J. C. Lovejoy, before the committee of the legislature of Mass., on the petition of Thomas H. Perkins and others, for the repeal of the liquor law, March 15th, 1853" et "Speech of Rev. J. C. Lovejoy, before the legislative temperance society, March, 30th, 1853." https://archive.org/details/cihm_52253 (consulté le 7/12/2019). pp.1-14.

Ratinaud, Pierre. (2018). « The brilliant friend(s) of Elena Ferrante: A lexicometrical comparison between Elena Ferrant's books and 39 contemporary Italian writers" in Tuzzi, Arjuna and Michele A. Cortelazzo. *Drawing Elena Ferrante's Profile: Workshop Proceedings, 7 September 2017.* Padova: Padova University Press. p.100.

Savoy, Jacques. (2012). « Attribution d'auteur : une approche basée sur l'allocation latente de Dirichlet. http://lexicometrica.univ-paris3.fr/jadt/jadt2012/Communications/Savoy,%20Jacques%20-%20Attribution%20d'auteur.pdf (consulté le 16/11/2019).

Tuzzi, Arjuna and Michele A. Cortelazzo (eds). (2018). "It takes many hands to draw Elena Ferrante's profile" in Tuzzi, Arjuna and Michele A. Cortelazzo. *Drawing Elena Ferrante's Profile: Workshop Proceedings, 7 September 2017.* Padova: Padova University Press. p.19.

Yellin, Jean Fagan, éd. (2008). *The Harriet Jacobs family papers*, Chapel Hill, University of North Carolina Press.

Zhao, Ying and Justin Zobel. (2005). "Effective and scalable authorship attribution using function words" in G.G. Lee et al, (eds). *AIRS*, 2005, LNCS 3689, p. 176. https://link.springer.com/chapter/10.1007/11562382_14 (téléchargé le 15/12/2019).