Roderick Random amidst Eighteenth-Century Fiction : A Computer-Aided Textual Analysis
Anne Bandry-Scubbi

Citer ce document / Cite this document :
doi : https://doi.org/10.3406/xvii.2009.2399
https://www.persee.fr/doc/xvii_0291-3798_2009_num_66_1_2399

Fichier pdf généré le 21/12/2018
RODERICK RANDOM AMIDST
EIGHTEENTH-CENTURY FICTION :
A COMPUTER-AIDED TEXTUAL ANALYSIS

Roderick Random is compared with nineteen other works of fiction published in the eighteenth century, organised in different corpora. With the help of data obtained by textual or linguistic analysis software, what critics have called “stylistic energy” is investigated by focussing on the use of superlatives, on syntax, the expression of time and the itemisation of the body, as well as on the manner in which the narrator is made to wield the first person and frequent substantives. Contrasting the relative frequencies of particular words or of word-lists related to a given feature thus reveals several lines of force of this novel often described as lacking structure.

Not many critical studies focus on Smollett’s way of writing, possibly because Damian Grant’s comprehensive 1977 *Tobias Smollett: A Study in Style* can be found in most university libraries, and also because interest in Smollett has waned since the early 1980s.1 In this paper, Philip Stevick’s 1967 “Stylistic Energy in the Early Smollett,” Paul-Gabriel Boucê’s chapter on “Aspects stylistiques”

1. It takes the Agrégation to spur interest on Smollett in France: *Roderick Random* was a set text in 1981 and *Humphry Clinker* in 1971 and 1992, as articles in *XVII-XVIII* show (see numbers 13 and 32).
(Romans 396-431) and Damian Grant’s 1982 “Roderick Random: Language as Projectile” are taken as starting points. Random is compared with nineteen other works of fiction published in the eighteenth century, organised in different corpora and analysed with the help of a computer, so as to draw out some of its characteristic features. My attempt at textual imaging aims at challenging the parallel drawn by Boucé between “corpus” and “corpse.”

I have built several corpora centred on Roderick Random for the purpose of this study. In Novels of the 1740s Jerry C. Beasley singles out Pamela, Joseph Andrews, Jonathan Wild, Clarissa, Roderick Random and Tom Jones as “the six major works” among “some three hundred works of fiction […] published between 1740 and 1749” (xii). Because Novels of the 1740s was published in the early 1980s, fiction written by women had not yet taken the place that it now has in the literary canon, partly to the detriment of Smollett. A first corpus, called “1740S,” consists of texts from the 1740s or early 1750s, all by different authors and of comparable length, with a balance between male and female authors, as well as between narratives about male or female main protagonists: Pamela; or Virtue Rewarded (Samuel Richardson, 1740), The History of the Adventures of Joseph Andrews (Henry Fielding, 1742), The Adventures of David Simple (Sarah Fielding, 1744), The Adventures of Roderick Random (Tobias Smollett, 1748), Memoirs of a Woman of Pleasure (John Cleland, 1748-49), The History of Miss Betsy Thoughtless (Eliza Haywood, 1751), and The Female Quixote or The Adventures of Arabella (Charlotte Lenox, 1752).

2. “Si tout le travail de dépouillement systématique des romans de Smollett selon les méthodes scientifiques de la linguistique moderne reste à faire […] la stylistique orientée vers la seule étude scientifique d’un texte a le tort de tuer toute vie, de traiter un roman comme un ‘corpus,’ cadavre littéraire figé dans la ‘rigor mortis’ ” (Boucé, Romans 429). The perceptive chapter on “Aspects stylistiques” which contains this statement is omitted from Boucé, Novels. Over the last four decades, tools for computer-aided textual analysis have evolved rather drastically, making it much easier to put into perspective the conclusions drawn from close readings of selected passages, not unlike the possibilities opened by medical imaging. The methodology of “samples,” as developed by Louis T. Milic in his seminal A Quantitative Approach to the Style of Jonathan Swift (1967) and Charles Muller in France, has been replaced by corpus analysis.

3. Beasley mentions Sarah Fielding and Eliza Haywood as authors whose works are not completely “deservedly neglected” (xii). The availability and quality of electronic editions also played a role in my selection of texts. See details in bibliography. No two programmes count “words” in exactly the same way, so the figures for the length of texts are approximations.
length of these texts varies between 98,000 words for *Memoirs of a Woman of Pleasure* and 224,000 words for *Pamela*, with *Roderick Random* at 200,000.

So as to take into account the constraints linked to a narrative in the first person, another large corpus, called 1stPERSON, is composed of fictional autobiographies: *The Life and Adventures of Robinson Crusoe*, *The Farther Adventures of Robinson Crusoe*, *The Fortunes and Misfortunes of Moll Flanders*, *The History and Remarkable Life of the Truly Honourable Col. Jacque commonly call'd Col. Jack*, *Roxana The Fortunate Mistress*, *Gulliver’s Travels*, *The Adventures of Roderick Random*, *Memoirs of a Woman of Pleasure* and *The Life and Opinions of Tristram Shandy*. The risk of bias caused by the prevalence of texts by Defoe (52% of the corpus) is balanced by the advantage of having three corpora of similar sizes including *Roderick Random*, the third comprising all of Smollett’s fiction. All the works apart from *The Adventures of Sir Lancelot Greaves*, *Travels through France and Italy* and *The History and Adventures of an Atom* are divided into two parts so as to have texts of comparable length. SMOLLETT is thus made up of 1,100,700 words over eleven parts, 1740S of 1,142,000 words from seven texts and 1stPERSON of 1,206,000 words from nine texts.

*Roderick Random* taken on its own makes up two different corpora: one in which it is divided into its 69 chapters and another where divisions follow the stages of the narrative.

Writing in the superlative?

Stevick examines “the stylistic energy which everyone recognises but no one describes” (712). Focussing on a passage from *Peregrine Pickle*, Chapter 24, he identifies the “peculiar combination of a brisk efficiency […] along with an ironic expansiveness and hyperbole” as “characteristic of […] nearly every paragraph in *Roderick Random*,

4. Some of these texts have already given rise to published and unpublished computer-aided textual analyses: see Bibliography under Bandry, and Bandry & Deconinck-Brossard.

5. Adapted from Boucé, *Novels* 104-24: Youth (Chapters 1-7), Journey (Chapters 8-13), London1 (Chapters 14-23), Navy (Chapters 24-37), Travels (Chapters 38-44), London2 (Chapters 45-52), Bath (Chapters 53-60), Jail (Chapters 61-64), and End (Chapters 65-69).
Peregrine Pickle and Ferdinand Count Fathom” (713). Boucè summarises this as “writing in the superlative” (Romans 422).

If one compares the use of superlatives in Random as a whole to that in other texts, however, it appears that Smollett actually uses fewer such forms than his contemporaries. Several estimates show that the proportion of superlative adjectives in Random is low – about half of what it is in the 1740S corpus as a whole –, and that the quantity of adjectives is slightly lower than average. In the chapters from Random in which more superlatives are used, the surplus is slight. Chapters 46 and 63 show the victimisation of Wagtail and Melopoyn. The banality of repeated superlatives (best, greatest, highest) is part of the game in which the victim is impervious to the irony of the language, even when, in the poet’s case, he tells his own story retrospectively. Intensifiers are not used more frequently in Random than in other texts either. Grant’s remark that “superlatives pursue each another” (Style 190) applies to the passage of Travels he focuses on but not to the text as a whole. This points to the limitations of considering word frequencies globally, as all three critics note that superlatives and intensifiers accumulate in the specific extracts they select. The method used here proves more fruitful for other stylistic phenomena.

While reading Roderick Random, one feels that one is being rushed through the text. The extensive paragraphs and long sentences are one reason. The inordinate length of the paragraphs can best be confirmed by just turning the pages of a printed copy. This is also true of sentences, of which an indicative measurement can be given with a caveat, as punctuation is often unreliable in electronic texts: an average sentence in Random contains 45 words (as in Moll Flanders and Roxana), while this mean value is 36 words for Gulliver’s

---

6. The list of words ending in -est, pruned of non-superlatives, added to occurrences of the most and the least are taken as a measure of the use of superlatives. Hyperbase, Concordance and Tree-Tagger have been used to compute this approximation. Methods will be detailed in “Les spécificités du roman de Smollett, Roderick Random” presented at the 10th International Conference on the Statistical Analysis of Textual Data (http://jadt2010.uniroma1.it/).

7. Boucè mentions “unsophisticated intensifiers” (Romans 401) or the abundance of all sorts of intensifiers (405) in the passages he selects. In Random as a whole, the relative frequency of the words he quotes is lower than or equal to their average value in the 1740S and 1STPERSON corpora.

8. The computer programmes at my disposal do make it possible to visualise concentrations of a given feature on a smaller scale than the divisions of the corpora.
Travels and 34 words for Joseph Andrews, 29 words for Pamela, but 69 words for Colonel Jack. Random’s long sentences are studded with parentheses, as the reader quickly notices: while the average is one open bracket for every 900 words in the 1740S corpus and one for every 730 words in 1stPERSON, Roderick Random has one for every 330 words (every 394 words in Tristram Shandy). Smollett later learnt to make shorter and more fluid sentences, as his revising of Peregrine Pickle for the 1758 edition shows (Boucé, Romans 393-94).

Another feature of this complex syntax is the heavy use of WH-relative clauses, much stronger in Smollett’s first three novels than in his later texts. Random stands out for this category of words both in the 1740S and in the 1stPERSON corpus, especially for which, who, whom and where: the average is one WH-word for every 45 words in 1740S and every 47 words in 1stPERSON, whereas Random has one in every 38 words (Joseph Andrews is just slightly higher at 36 words). Whereby and whereupon are also specific to Smollett’s intricate syntax, although far less frequent (respectively 10 of the 12 and 24 of the 25 occurrences in 1740S belong to Random).

The regular distribution of WH-words taken globally over Random shows that their use is a stylistic trait of the early Smollett, not linked to subject matter or context.

The quantitative importance of WH-words compensates for the unexceptional frequency of and: one occurrence for every 33 words on average, the same as the mean value in the 1740S corpus. Smollett seems to find his tempo for this very common word from the episode of the Thunder onward (Chapter 24), with a rate of 3.17% of all occurrences, which he maintains in all the texts considered in the SMOLLETT corpus.

11. The null hypothesis for statistical tests about a text is that all words are distributed equally throughout the corpus. Specific words are words “abnormally” frequent in one part of the corpus (Lebart & Salem 260).

12. Sterne also takes a while to find his tempo, with fewer dashes in the first volume of Tristram Shandy than in the rest of the work, as shown in Bandry, Shandy 47.
syntax, as is shown by the low rate of WH-clauses (1.72%) and a strong use of and (slightly over 4%, or one occurrence for every 24 words). In Random, the two features reinforce one another, whereas one or the other predominates in all the other 1740S texts, except for Betsy Thoughtless in which both are low (fig. 1).

When the frequencies of WH-words and of and are taken together, Random ranks second behind Pamela. To Richardson’s (and to a lesser extent Defoe’s) practice of juxtaposition, Smollett prefers complex relative and interpolated clauses. His fondness for asyndeton, accumulations, lists of verbs, adjectives or substantives separated by commas, in which “redundant detail serves as a comic surcharge,” also accounts for his average use of and despite the luxuriance of the text and “the crush of words” (Grant, Style 141, 135). The abundance and variety of the vocabulary appear through the high quantity of hapax legomena used in Random: in the 1740S corpus, 28% of the words used only once belong to Smollett’s novel, rather than its expected share of 17%. It comes second to Memoirs of a Woman of Pleasure.15

---

13. Richardson to Lady Bradshaigh, 14 February 1754, Selected Letters 289.

14. Smollett also likes nominalisations, as surpluses for words ending in -ment, -ity, -ness and particularly -tion show. The names Potion and Lavement only account for 14 out of 2,138 words ending in -ion and 15 out of 501 ending in -ment.

15. The expected rate of a given feature is the proportion of the corpus taken up by the text: Roderick Random makes up 17% of 1740S, so a deviation occurs when the actual occurrences of a word in Random total more or less than 17% of the occurrences of the corpus. A deviation is statistically significant if it exceeds a threshold often estimated at 5%, as is apparent in Hyperbase diagrams. Cleland’s pornographic text makes up 8% of the 1740S corpus but 16% of the hapaxes of the corpus belong to it. In the 1stPERSON corpus, Shandy comes before Woman of Pleasure and Random.
Time is collapsed

The second feature which causes the sense of being rushed through the text, according to Stevick, is that “time is collapsed” (714). The adverb immediately belongs to Random’s specific words in the 1740S corpus: it appears almost twice as often as expected, with 170 of the 530 occurrences (32% rather than the expected 17%). Immediate is also more frequent than in the other texts (28% of all occurrences), along with phrases noted by Stevick and Boucé: no sooner than (123 in Random, 50% of the occurrences in the corpus), on/of a sudden (13/38), in a moment (16/32), in a trice (5/12), in a twinkling (6/6), on the spot (4/7), upon which (116/167), whereupon (24/25).

A more extensive quantitative analysis shows that Random contains twice as many verbs, adjectives, adverbs and phrases expressing speed and precipitation as expected: 35% rather than 17%. It is moreover the only text with a marked positive deviation (fig. 2). By contrast, Pamela is decidedly slow (5% of these words, instead of the expected 20%), while speed is not an issue in the other novels.
The terms taken into account are the forms of *hasten*, *hurry*, *immediate*, *precipit-*,-, *quick*, *rapid*, *rush*, *speed*, *sudden*, *swift*, *veloc-*.

Although words beginning with *precipit-* are not very frequent, they stand out in *Random*: it contains 50% of the 34 instances in the corpus and the greatest variety with occurrences of *precipitation*, *precipitately* and *precipitate* in its different forms. The 170 uses of *immediately* combined with the accumulation of other expressions of speed (141 occurrences) greatly contribute to the impression of being rushed through *Random*.

**Figure 2**

Indications of time in general are also strong in *Random*. Although the frequency of time adverbials is slightly lower than

---

16. The list was constituted from the frequency lists of the corpus and of the individual texts, as well as from *The Oxford Thesaurus*, with the meaning checked from context as displayed in concordances. Three words were eliminated: *dispatch* (mostly used in the sense of *send*), *expedit-* (most often *expedition*, in the sense of *travels*) and *fast* (mainly in the phrase *fast asleep* or meaning *firmly*). In “Language as Projectile,” Grant provides a list of 66 verbs which launch “oaths, curses, ‘miserable expletives’ and ‘opprobrious epithets’ ” (144).

RSÉAA XVII-XVIII 66 (2009)
average in the 1740S corpus, indications of duration stand out.\textsuperscript{17} \textit{during} and \textit{while} are used significantly more than in other texts, along with \textit{whilst}, \textit{until} and \textit{till}.\textsuperscript{18} \textit{Pamela} is again the text in which this feature is the lowest, while \textit{Woman of Pleasure} ranks second. If one considers precise indications of duration (\textit{year}, \textit{month}, \textit{week}, \textit{day}, \textit{hour}, \textit{minute}, \textit{seconds} and their derivatives), \textit{Random} comes first for all except \textit{year}, with a cumulated share of 25\% in 1740S.\textsuperscript{19} \textit{Day} is particularly high (it is used twice as often as in \textit{Joseph Andrews}) and greatly helps create the impression of a quick succession of events as it most often follows \textit{next} (69 times). The times of day are also mentioned more frequently in \textit{Random} than in others texts, especially \textit{night}, and to a lesser extent \textit{morning} and \textit{afternoon} (rather than \textit{evening}); it ranks second after \textit{Pamela} for the mention of particular hours (\textit{o\'clock}). As for relative time, \textit{Random} stands out for its use of \textit{before} and \textit{after}, rather than for \textit{now}, strong in both \textit{Pamela} and \textit{Woman of Pleasure} (for very similar reasons): contrary to these heroines, Roderick as a narrator focuses more on the succession of actions than on their immediate consequences.

By making the first-person narrator mention time often and in several different ways, more than in contemporary texts by other writers, Smollett makes the story proceed at a brisk pace. A similar insistence is put on the body.

\textbf{Body language}

One way to position a text with respect to a corpus is to examine a list of words and their frequencies related to a given feature: relative frequencies of particular words or of the list as a whole can be compared. Taken globally, such a list of terms designating parts of the body shows that in the 1740S corpus, \textit{Random} comes second for its positive deviation from the average use after (unsurprisingly)

\endnote{17}{Biber provides the following list of time adverbials: \textit{afterwards}, \textit{again}, \textit{earlier}, \textit{early}, \textit{immediately}, \textit{presently}, \textit{soon}, \textit{yesterday}.}
\endnote{18}{23\% of the occurrences of these words are in \textit{Random}, rather than the expected 17\%.
\endnote{19}{Days are referred to 307 times in \textit{Random}, hours on 107 occasions, years on 73, weeks on 63, months and minutes on 50. \textit{Minute} is a time unit in 53 of its 61 occurrences in 1740S, as a concordance shows (13 occurrences in \textit{Random}). Only Pamela details the days of the week, when she dates entries in her letters and her diary.}
Memoirs of a Woman of Pleasure.\textsuperscript{20} Novels written by women (David Simple, Betsy Thoughtless and The Female Quixote) contrast with these two texts by referring to the body far less often. In the 1stPERSON corpus, Random comes after Tristram Shandy, Memoirs of a Woman of Pleasure and Gulliver, while all the texts below average are Defoe’s. The ratio of body words to total occurrences puts Random among the texts which mention the body to a (statistically) significant extent (fig. 3):

\begin{center}
\includegraphics[width=\textwidth]{figure3.png}
\end{center}

If occurrences of basic verbs and organs are taken as an indication of the degree to which the senses are used, differences between texts are slight when all senses are considered together. The small deficit of Betsy Thoughtless matches Pamela’s surplus (one fifth more than expected, coming from sight). However, as Grant notes, “part of Smollett’s comic technique [is] to describe the operation of each sense in isolation” (Style 115). Fewer eyes and forms of see in Random combine with the importance of night and darkness in general to neutralise sight to a certain extent and foreground the other senses, usually less resorted to in literary texts. Random quite expectedly stands out for a strong use of smell, with 58\% of the 86 occurrences of nose/s and the forms of smell in the 1740S corpus. Although the ratio for hear (lemmatised) and ear/s is slightly under average (but not in a statistically significant way), occurrences of noise/s, sound/s and voice/s show the importance of hearing in

\textsuperscript{20}. Woman of Pleasure has twice as many terms as expected if 1740S is taken a norm; Random has a third more. The list was compiled from the frequency lists of the corpora and of the individual texts.
Random; with 27% of these words, it contrasts vividly with Pamela’s 7%, a strong deviation from its expected 20%. Taste as a sense cannot be indicated in this manner because the abstract meaning prevails, as a concordance shows. Neither can such a crude estimation do justice to the great variety of words used for physical contact in Random and in Woman of Pleasure. The verb “touch” does appear far more often in Cleland’s text than in the rest of the 1740S corpus (43% of 182 occurrences). The most frequent body words help refine this approach (fig. 4, the ratios indicate the proportion of each word to the total number of occurrences in the text or corpus).

<table>
<thead>
<tr>
<th>1stPERSON</th>
<th>1740S</th>
<th>RANDOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 hand</td>
<td>0.082%</td>
<td>heart</td>
</tr>
<tr>
<td>2 head</td>
<td>0.064%</td>
<td>hand</td>
</tr>
<tr>
<td>3 hands</td>
<td>0.052%</td>
<td>eyes</td>
</tr>
<tr>
<td>4 heart</td>
<td>0.045%</td>
<td>head</td>
</tr>
<tr>
<td>5 eyes</td>
<td>0.034%</td>
<td>hands</td>
</tr>
<tr>
<td>6 face</td>
<td>0.031%</td>
<td>face</td>
</tr>
</tbody>
</table>

In Random, hand comes first (like in Shandy, and by way of consequence in the 1stPERSON corpus), while heart heads the 1740S list and Pamela (it comes fourth in Random, with half the relative frequency of Richardson’s text but ratio similar to that in other 1740S texts; Defoe and Swift rarely refer to it). Eyes predominate in Memoirs of a Woman of Pleasure, as voyeuristic pleasure combines with sensual contact (skin and thighs run high and so do hand/s, of course), even if the body is first appreciated globally as a figure (a quarter of the occurrences in the 1740S corpus, and one fifth in The Female Quixote). If hand and hands are joined, they show up in similar rates in Random, Woman of Pleasure, and Pamela, but all three texts give precedence to Shandy. While Smollett tells more of fisticuffs, Sterne uses hand/s for sentimentalism and nose/s for bawdiness (with three fourths of 1stPERSON nose/s, leaving 14% for

21. This aspect will be developed under the aegis of LAPASEC (Landau Paris Symposium on the Eighteenth Century) with a study of “Changing Taste in the Use of Touch: from Haywood to Burney.”

22. A concordance shows that figure is far more often an indication of the body than a verb.
Random, for smell and fights). Heads are more often struck or butted than shaken in Smollett’s text (as a concordance shows) and fewer of them belong to Roderick than to other male characters. Comparison between the texts evidences that Smollett breaks the body down into its different parts more than other writers, which efficiently creates Random’s atmosphere of brutal encounters.

If one focuses on the different sections of Random, it comes as no surprise that the episode in the navy (chapters 24 to 37) should contain the most body terms (fig. 5). This part makes up 17% of the corpus but 21.5% of body terms belong to it. To an even greater extent than in the book as a whole, the accumulation of physical references prevails over the use of any particular word from the list. Contrary to expectations, the jail chapters (61-64) have the fewest body words, less than half the expected ratio (3% for 7% of Random). Figure and beard, the most specific body terms in this section, help sketch Melopoyn, the pitiable consequence of too great a concentration on the preoccupations of the mind. Without a Strap to shave/save him, he belongs to the category of bearded madmen, like the owner of Ralpho (chapter 13) or the wild Rourk Oregan (chapter 49). By making his hero distance himself from his “strong passion for the Belle Lettres [sic]” through the poet’s “grotesque figure” (Chapter 61), Smollett brings him closer to his final destination: “Having exposed himself to most of the sensations that the capital city, the British navy, and the colonies have to offer, Roderick […] establishes himself on a country estate […] a place where the hero’s body is safe from overwhelming sensation” (Douglas 48). While hand and head characterise the youth, journey and navy episodes, heart and eyes come later, in Bath and in the final part: Roderick must learn to use the senses which prevail in texts about women in order to deserve his Narcissa.

Body parts and characters are often related to the narrator through the use of my, which is Random’s most specific word in the 1740S corpus.

---

23. For Sterne’s use of hands, see Bandry, “Les livres de Sterne.”
24. Face, interestingly, occurs at a comparable rate in all texts.
25. As shown by a factor analysis (provided by Hyperbase).
The first person or “principal personage”

A concordance makes it possible to analyse what my refers to, bearing in mind that Roderick is not the only one who uses the first person. Body parts preceded by my are first abstract, with my heart (47), my soul, not very different from my life (37), then more concrete with my hand (30), my face (21), my blood (20), eyes, head, breast, back, bosom, neck, senses, tongue (16 to 8 of each), arms (7, differentiated from occurrences meaning weapons), and more dramatically, my wounds (9). Most bodies are itemised: eyes and hand(s) belong more often to another male character (not always Strap) than to Roderick but are directed to or against him. As Douglas notes, in this novel “the body […] only learns to [function in society] after shocks and pains” (49).

These are assessed by the second highest category of words preceded by my, which designate the situation or state in which the speaker finds himself (or herself): my passion (28 times), my situation (27 times), my fortune (26 times) or my misfortune (18 times), my business (21 times), my appearance (18), my condition (16), my fate (14), and my happiness (13) are the most frequent occurrences. Quite
logically in a first-person narrative, these words apply more to the speaker than to any other character, as the determiners show.

The physical and psychological interactions between characters make up the story, in which Roderick is no loner. The highest number of words with my as a determiner designate people: my uncle (79 times), my friend (69 times) and my friends (18 times), my father (68 times out of 125 occurrences of the substantive, although Roderick’s father disappears after the first chapter to pop up again in chapter 66; Miss Williams uses my father 10 times in chapter 22), my grandfather (32 times out of 36 occurrences of grandfather, duly concentrated in the first six chapters), my companion (32 times), my acquaintance (15 times) as well as my dear (29 times), but also my rival (11 times). My mistress ranks first in the female designations (50 times out of the 79 uses of the substantive) but serves for several statuses and people: Narcissa’s aunt, Narcissa herself, but also Melinda. My lady, my mother, my charmer, my wife, my female cousins all appear far less frequently, with 9 to 7 occurrences each. Like Tristram Shandy (the second highest user of my in the 1st PERSON corpus), if less comfortably so, Roderick’s world is peopled with characters which, as a first-person narrator, he relates to himself. The fact that his and with are even more specific of Random than my in the 1st PERSON corpus confirms that the hero is “surrounded with bodies” (Douglas’s title for her chapter on Smollett’s first novel).

Although the global amount of personal pronouns and their obliques in Random conforms to the average in both non-smollettian corpora, I is used far less than in Defoe’s fictional autobiographies and in Richardson’s epistolary fiction: Pamela, Robinson and Moll appear much more self-centred than Roderick, when one realises that their use of I makes up 4% of all the words of their stories (4.6% for

26. His father appears 12 times and her father 11 times, referring to various families: 109 uses of father in all, of which slightly more than half are preceded by my.

27. I take the term obliques from Freedman, which she uses to designate “thee, thy, thine” and “ye, your, yours” in relation to thou and you (2). Burrows uses the category of “pronouns” for we, our, and us (34). Such words make up 13.5% of all occurrences, which is the average value in both non-smollettian corpora (1740S and 1stPERSON), discounting Pamela and its particularly high rate of 18%. 1740S: 14% with Pamela, 13% without. 1stPERSON: 13.5%. This rate is much lower in Smollett’s Atom and Travels (7%), which points to a link between genre and the use of such terms. Smollett never again used as many as in Random.
Pamela, but the letter-form probably plays a role), whereas it constitutes only 2.5% of those used by Roderick. Gulliver is made to use I at a comparable rate (2.7%), while Fanny Hill and Tristram Shandy resort to it less often; it makes up respectively 1.78% and 1.52% of the words they use to tell their stories. Such a ratio is not easy to spot, as imitations of Sterne show: spurious Tristrams used I too often, while spurious Yoricks did not use it enough for their continuations of A Sentimental Journey to read like genuine Sterne texts.

Roderick is not a loner, as was pointed out earlier. The use of the first person plural in Random is slightly above average for the 1stPERSON corpus, with 0.9% of total occurrences, a rate comparable to Colonel Jack (1.1%) and Robinson Crusoe (0.8%, the mean value). In The Further Adventures, Robinson organises a society of humans around him, rather than of objects and animals as in The Life and Adventures, which entails a strong use of the first person plural (2% of all occurrences) rather than of the singular. In Moll Flanders and Roxana, this rate is much lower (respectively 0.6% and 0.5%, 0.4% for Pamela): male characters of fictional autobiographies seem more sociable than female ones. To some degree, the distribution within the chapters of Random shows an alternation between the first person singular and plural, my and our, I and we (Figure 6). Random does not stand out for its global use of the first person in the 1stPERSON corpus, with neither positive nor negative deviation, but the detailed analysis shows that, compared to other first-person narrators, Roderick is less focussed on his person than on the world as it relates to him.

28. In “On peut compter sur Moll” we had noted the unusual presence of I in the first position of the frequency list, with twice as many occurrences as the, usually the most frequent word. I comes at a more usual fifth rank for Random, with the customary the, of, to, and and before it.
30. We, us, our, ours, and ourselves.
In his fine study on “Language, Structure and Vision” in Roderick Random, Ian Campbell Ross shows how “the moral disorder perceived by Smollett is consistently and coherently created by verbal repetition” of terms expressing the concepts of inhumanity, injustice and dishonesty (53). Yet, if “the language of disorder provides a solid framework for otherwise separate and causally unrelated episodes” (63), other lines of force have emerged, by contrast with other texts: speed and precipitation, language of the body, relating the (admittedly chaotic) world to the narrator. Substantives used very frequently provide another. One of Roderick’s constant preoccupations is to be the gentleman he was born to be. Contrary to Moll Flanders who is

31. Françoise Deconinck-Brossard and I have often found that the most frequent non-grammatical words are particularly significant for the logic of the story: in Random these are man, day, house, fortune, friend, manner, person, gentleman, night, and money. Working with large corpora adds the possibility of identifying words which are frequent and “specific” to a given section or text.
made to overuse the word gentlewoman in a skewed meaning when very young and then wield it with great irony as a retrospective narrator. Smollett’s character refers to this status regularly throughout the text (141 occurrences of gentleman). Combined to his constant concern with situation (109), behaviour (119), fortune (171), appearance (92) and expence (67), Random’s most specific substantives in the 1stPERSON corpus, such a preoccupation of the hero with his social status provides a powerful uniting principle, part of the “uniform plan” organised around the “principal personage” (Smollett, Fathom 43). His final poise results from the tension between the centrifugal picaresque force and the centripetal energy of the tennis-ball from his mother’s dream.

The data examined here often shows Roderick Random to be close to Woman of Pleasure. Although the stamina of these texts and their heroes is geared towards different purposes, the stylistic energy which animates both texts is high. It is hoped that the analysis of textual imaging proposed in this paper provides an insight into the workings of the internal anatomy of The Adventures of Roderick Random.

Anne BANDRY-SCUBBI

Université de Strasbourg

32. Half of the 70 occurrences of gentlewoman are clustered over a few pages (Bandry & Deconinck-Brossard, Moll 182).

33. Situation, expence, and fortune are also Random’s highest specific substantives in the 1740S corpus, with appearance not far behind.

34. I am very grateful to Damian Grant for Random discussions before and during the writing of this paper.
LIST OF WORKS CITED

PRIMARY SOURCES


SECONDARY SOURCES


SOFTWARE


