

## Key Pronouns through Wmatrix in a Novel of Formation: Conrad’s *The Shadow-Line*

Giuseppina Balossi

Liceo Scientifico e Musicale ‘G.B. Grassi’, Lecco, Italy  
[g.balossi@libero.it](mailto:g.balossi@libero.it)

### Abstract

The advent of the digital age has had an enormous impact on the way we research, teach and think about language ([23]:1). In digital humanities, the application of computer-assisted methods can facilitate investigation of corpora, (*i.e.* digitised samples of language in use), lead to discoveries barely detectable with the naked eye and help put interpretation by intuition to test. Such methods can also assist in investigation of the language of single texts and make our close reading more effective. This article aims to suggest how we may investigate the narrative voices in a work of fiction through the program Wmatrix ([30]). The case study under analysis is Conrad’s *Bildungsroman The Shadow-Line, A Confession* (1917), a story revolving around a young inexperienced sea-captain who, during his first command of a ship, has to overcome a series of difficulties to accomplish his mission. In this work, the neat division between the *I* protagonist-narrator’s internal world and the external adult world he has to confront lends itself to investigation of the first person and other personal pronouns in the whole work. Given that stylistic analysis of literary texts is fundamentally a comparative process, I am here interested in comparing the target text, *The Shadow-Line* to two comparison texts by the same author. The keyness statistics for the pronouns and their detailed analysis through concordances will contribute to placing the *I*-voice at the centre of the narration and to identifying the foregrounded patterns of pronoun use that convey the *I*-voice and the other narrative voices<sup>1</sup> throughout the story.

---

1 By *I*-voice or narrative voice, I refer to all the pronominal forms (*e.g.* *I*, *me*, *my*, *mine* and *myself*) and to their grammatical functions (*e.g.* subject, object, *etc.*).

L'avvento dell'era digitale ha avuto un enorme impatto sul modo con cui noi facciamo ricerca, insegniamo e pensiamo il nostro linguaggio ([23]: 1). L'applicazione di metodi computazionali agli studi umanistici rende fattibile l'analisi di corpora (campioni digitalizzati della lingua in uso), porta a scoperte impossibili con altri strumenti e offre elementi di verifica oggettivi all'interpretazione tradizionale dei testi. Essa aiuta poi efficacemente nell'analisi del linguaggio di singoli testi e rende la lettura ravvicinata più efficace. Questo articolo ha lo scopo di suggerire come si possono analizzare con il software Wmatrix ([30]) le voci narranti in un'opera narrativa. Oggetto di studio di questo articolo è il *Bildungsroman The Shadow-Line, A Confession* (1917) di Conrad, un racconto che ha come protagonista l'io narrante di un giovane e inesperto capitano di marina che durante il suo primo comando deve superare una serie di difficoltà per portare a termine la sua missione. In quest'opera la netta divisione tra il mondo interiore dell'io-protagonista e il mondo esterno con cui si deve confrontare si presta all'analisi dei pronomi di prima persona e degli altri pronomi personali nell'intera opera. Considerato che l'analisi stilistica di testi letterari implica un processo fondamentalmente di comparazione, si confronta il testo target *The Shadow-Line* con due romanzi simili dello stesso autore per analizzare le categorie grammaticali predominanti dei pronomi. I risultati statistici dei pronomi e la loro analisi tramite concordanze contribuiranno a posizionare l'io narrativo al centro della narrazione e a identificare l'interazione tra l'io narrante e le altre voci narrative.

## 1. Introduction

By definition, the term 'corpus-assisted methods' involves the analysis of large corpora. Biber ([7]: 15) defines a corpus as "a large collection of principled texts stored on computer [...] designed to represent a textual domain in a language". However, Biber did not exclude the applicability of corpus linguistics also to literary texts, thus infringing the basic assumptions of the corpus approach concerning size and representativeness ([9]). Indeed, within the field of literary stylistics we come across plentiful corpus-based studies that are focussed on an individual target text of a given author, compared against a reference corpus of the same author. For example, Stubbs ([37]: 5-24) carried out a quantitative analysis of the frequency distribution of single words and the lexico-grammatical patterns in Conrad's *Heart of Darkness* with one of the first versions of WordSmith Tools ([34])<sup>2</sup>, and pointed out that collocations help disclose connotations and sometimes the idiosyncratic meanings of words. McIntyre and Archer ([26]) employed the UCREL Semantic Annotation System (*cf.* Section 3) to investigate the distribution of semantic domains in Alan Bennett's play *The Lady in the Van* in order to show the potentialities that a quantitative approach brings to the study of a character's mind style, a phenomenon which has traditionally been studied qualitatively. A further example of corpus-based methods implemented on a single target text is found in my own study *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's The Waves* ([5]). I here investigated quantitatively and qualitatively the claim, subject to much debate in part of the literary arena,

---

2 The latest updated version of WordSmith was issued in 2020.

whether the language of each of the six characters can be distinguished from each other. This broad issue is tested by means of a series of statistical comparisons of each character's part-of-speech and semantic domains through the program Wmatrix ([30]).

In digital humanities ([32]), corpus-aided methods are largely employed in the study of language from different perspectives, amongst which are intra-textual "variations in the style of a single author or text, including chronological change, the dialogue or narration of multiple characters or narrators in a single novel, and other perceived shifts in style" ([21]: 17ii); such methods are also increasingly being applied in pedagogical contexts (see, for example, [26]: 253-270) to look at the stylistic properties of literary texts arising from patterns of use of linguistic features.

The stylistic properties investigated in this study are those involving pronouns, a grammatical category that falls into the larger category of function words, which has traditionally been the object of quantitative studies carried out in the field of authorship attribution ([15]: 271-280; [22]: 59-66), investigations of literary idiolects, and stylistic changes in a given author's artistic production ([10]; [38]). More recently, the use of pronouns and their role in narratives have been investigated in corpus stylistics ([25]; [4]; [27])<sup>3</sup> to study their stylistic features and effects on the reader since pronouns

delineate a narrator's subject position in relation to objects/others. In effect the pronoun simultaneously determines, designates, identifies, refers to and (re-) affirms a particular narratorial role [...]. A narratorial voice often works as a focalizing 'window'- a particular, positioned perspective on fictional worlds. ([19]: 2)

In the present analysis, I will investigate the subject position of the *I*-protagonist and narrator in Conrad's *The Shadow Line*, in relation to the other pronouns within the same text. With the aid of Wmatrix tools for POS analysis ([30]), the distributions of personal pronouns in the target text can be readily determined, as well as any statistically significant ('key') differences in pronoun frequency between the target text and similar texts by the same author. Meanwhile, the concordance lines generated for the pronouns with the highest frequency will enable observation of the *I*-narrator's subject position and its dynamics of interaction versus the other participants deployed at key moments in the narrative.

## **2. The Shadow-Line: the target text compared with Heart of Darkness and The Secret Sharer**

*The Shadow-Line, A Confession* ([14], hereafter TSL) is the story of an unnamed young man who, after deciding to leave his position as a mariner on a steamship and return to England for good,

---

3 This article builds upon the paper I presented at the Annual Conference of the Poetics and Linguistics Association (published online in PALA Proceedings, [4]), and at AIUCD Conference 2019.

is unexpectedly appointed captain of a ship. The protagonist, now older, recalls the sea voyage from Bangkok to Singapore, and the obstacles he had to face. This *Bildungsroman* or novel of formation ([20]: 293-311) takes on the structure of a Proppian initiation ritual<sup>4</sup> in which the challenge facing the young captain involves his ability to sail the ship and manage the crew, while encountering severe natural obstacles accompanied by an obscure and menacing adversary (the dead former captain). However, Ransome (the helper) is always on his side from the beginning to the end of his ordeal. In the end the young captain will manage to bring his ship safely back to harbour.

The grammatical usage of pronouns in TSL, as the target text, will be compared quantitatively to that in *Heart of Darkness* ([13], HD) and *The Secret Sharer* ([12], TSS). These two reference texts have been selected since all three works belong to the genre of the novel of formation, and all three are first-person narratives (Matlow in HD, and an unnamed captain in his first command in both TSS and TSL), each recalling in retrospect his first sea voyage as a young captain, the obstacles he had to face and the struggles within himself and with the adult world that challenge his initial naïve, idealised worldview and lead to his maturing process.

### 3. Methods: text digitisation and POS analysis in Wmatrix

The analysis of the personal pronouns in Conrad's *The Shadow Line* was carried out using Wmatrix ([30]), a web-based environment containing corpus annotation tools for grammatical and semantic analysis developed by UCREL<sup>5</sup> at Lancaster University.<sup>6</sup> Similarly to other text analysis software, such as WordSmith Tools ([34]), Antconc ([2]), Wmatrix can generate frequency lists, key word lists, n-grams/clusters, collocates and word clouds that display the more frequently occurring words in a larger font for any text uploaded into the system (Antconc does not have this facility). The added value of Wmatrix is that it can perform automatic semantic analysis, which is beyond the scope of the present analysis ([30];[3]).

To process an automatic analysis in Wmatrix, users need to upload their texts into a web browser, then the Wmatrix Tag Wizard automatically carries out POS (part-of-speech) and USAS (UCREL Semantic Analysis System) tagging. The frequency and relative frequency, grammatical or semantic tags are generated and can be sorted out either alphabetically or in order of frequency; examples of each of these can also be viewed in context in concordance lines. Part-of-speech tagging is processed by CLAWS (Constituent Likelihood Automatic Word-tagging System

---

4 An initiation ritual marks the entrance of a youth into adulthood, and the concept is used metaphorically by Propp to categorise a narrative stage performing the same function. Propp applied this notion fully to his structuralist analysis of the tale, especially in his work *The Morphology of the Folktale* (1968).

5 University Centre for Computer Corpus Research on Language.

6 The new version Wmatrix4, which replaced Wmatrix3, is here employed. The introduction to Wmatrix and its applications can be accessed at <http://ucrel.lancs.ac.uk/wmatrix/>.

([17];[18]), a tool designed to assign grammatical tags to the words in a text. The first step that the CLAWS tagger carries out is the tokenization of a text and its morphological and grammatical analysis, then it allocates words to their parts-of-speech on a non-contextual basis, assigns each word to its grammatical category and in a subsequent phase, establishes its most probable word-tag taking into account the linguistic context in which the word occurs ([30]: 63-65). The accuracy rating of CLAWS is between 96-7%; if users want to increase the accuracy level they can manually intervene on the processed data. Once a target text is processed, we can generate key words and key semantic domain profiles by comparing a target corpus against other reference corpora stored in Wmatrix (e.g. samples from the BNC) or uploaded by users in the Wmatrix environment.

For a keyness analysis, two word-frequency lists, one from the target corpus and the other from the reference corpus are compared. Keyness is not measured in terms of relative frequency alone, but using a statistical measure called Log-likelihood (LL hereafter). This statistical test establishes how significant the relative frequencies are between two corpora, or texts, taking into account their respective sizes. The normal level for statistical significance is the 95% confidence level ( $p < 0.05$ ) Chi-square value, which equates to an LL value of 3.8 or above. The recommended LL values for a word/tag to be statistically significant is above 6.63, as this is the cut-off point representing 99% confidence in its significance ([29]: 519-549); the higher the LL value the more significant is the difference between the two corpora. Once a keyness analysis is performed we can see the list of the overused word/tags+ or underused words/tags in the target corpus relative to the reference corpus sorted on their log-likelihood values<sup>7</sup>. Here, *overuse* means 'used with significantly higher frequency in the target than the comparison corpus', and *underuse* means 'used with significantly lower frequency in the target than the comparison corpus'.

Most studies that perform keyness analysis focus on the lexical level only, whereas the present study is one of the few to extend the concept to the level of grammatical word class (see [29];[5]). From a purely quantitative perspective, keyness analysis helps identify keywords ([35]), key POS and key semantic tags in a corpus, while from a qualitative perspective overused or underused words/tags indicate the distinctive features of a given text that are either over-represented or under-represented as compared with a reference corpus ([24]). For sake of clarity, I here reproduce in Figure 1, an extract from a table showing the first four hits obtained after carrying out a keyness analysis of parts of speech (POS) between TSL (target text) and one of the reference texts (TSS), where O1 is the observed frequency in the target text and O2 in the reference text sorted on the item's LL value. If we look at the PPIO1tag (1st person sing. objective personal pronoun [me]) we note that in O1 this category has a higher number of occurrence (372, equal to a percentage value of 0.99%) and an LL value (19.17) indicating overuse. On the left of the

---

7 The LL formula can be worked out using Rayson's LL calculator at: <http://ucrel.lancs.ac.uk/llwizard.html>. The Chi-square test could also have been used as it gives similar results to LL. Practitioners of statistical tests may want to look at two enlightening articles written by Dunning (*Errorc. L'origine riferimento non è stata trovata.*: 61-74) and Rayson et al. (*Errorc. L'origine riferimento non è stata trovata.*: 1-12) about the different applications and issues regarding statistical significance tests such as LL and Chi-square tests.

table, we can click on ‘List1’ or ‘List2’ to see the pronoun words comprised in PPIO1, or click on the concordance links if we want see keyword context results for the PPIO1 tag.<sup>8</sup> The word cloud displayed below the table shows the most frequently occurring tags in a larger font, amongst which is the PPIO1 tag.

	Item	O1	%1	O2	%2	LL
1 List1   List2   Concordance	NNB	186	0.49	1	0.01 +	114.61
2 List1   List2   Concordance	PPIO1	372	0.99	90	0.60 +	19.71
3 List1   List2   Concordance	CST	299	0.79	80	0.53 +	10.66
4 List1   List2   Concordance	JJ	2672	7.09	951	6.34 +	8.79

Figure 1: Sample of a part-of-speech keyness analysis, with its associated word cloud.



Figure 2: Key POS cloud.

The keyness analysis of the POS frequencies of the personal pronouns in TSL is produced through comparison between the target text (TSL) and the reference texts (IID and TSS) respectively. The digitised versions of the texts utilised for quantitative analysis are from Project Gutenberg eBooks (<https://www.gutenberg.org/>). The e-texts come in plain text format (TXT), thus they were already perfectly suitable to be fed into and processed by Wmatrix. However, before loading each text into the system, angle brackets were inserted around the parts of the text I wanted the program to ignore (*e.g.* the cover page and the chapter numbers). The next step entailed selecting three different sample texts I would use for qualitative analysis. The computer-based analysis of the personal pronouns in TSL follows the stages indicated below:

1. Corpus design:

-choice of target text and reference texts (*cf.* Section 2).

2. Pronoun retrieval, through the CLAWS POS (part-of-speech) tagger, built into Wmatrix.

- keyness lists, listing the pronoun tags in order of keyness.

- concordances, showing the occurrence of the personal pronouns in their co-text.

<sup>8</sup> In Figure 1, the Log Ratio is not displayed next to the LL column. This statistical test was added to Wmatrix 4, the Log Ratio statistic is an effect-size statistic, not a significance statistic since it represents how big the difference between two corpora are for a keyword (for further details, see: <http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/>).

- frequency lists, listing the pronouns in order of frequency.<sup>9</sup>

3. Interpreting the data:

-compared to the three steps in Stage 2 which are quantitatively based, this stage of the process, although it makes use of the 'concordance' option in Wmatrix, is qualitative.

#### 4. Results: key POS tags, and key pronouns

Table 1 shows the statistically significant personal pronoun tags<sup>10</sup> sorted in descending order of LL values produced by Wmatrix, when the target corpus *The Shadow Line* (TSL; corpus size 37,710) is compared with the first reference text *Heart of Darkness* (HD, corpus size 37,661) and the second reference text *The Secret Sharer* (TSS, corpus size 14,997).

<i>TSL compared with HD</i>	<i>LL/+</i>	<i>TSL compared with TSS</i>	<i>LL/+</i>
APPGE	28.89	PPIO1	19.71
PPIO1	24.81	PPHS1	8.24
PPHS1	18.71		
PPIS1	18.56		

Table 1: The most 'key' POS tags in TSL

Key:

APPGE - possessive pronoun, pre-nominal (e.g. my, your, our)

PPIO1 - 1st person sing. objective personal pronoun (me)

PPHS1 - 3rd person sing. subjective personal pronoun (he, she)

PPIS1 - 1st person sing. subjective personal pronoun (I)

More robust results are recorded when the target text is compared with HD than with TSS, producing both a higher number of significantly overused POS tags (*cf.* Appendix 1), and higher LL values. This is predictable given the evident similarities between TSL and TSS. In this regard, Benson's ([6]: 46-56) study of *Conrad's Two Stories of Initiation* defines the two novellas as "twin stories of initiation into maturity" ([6]: 46) and states that "there is in *The Shadow Line* a figure

9 For reason of space, the concordances for the occurring pronouns, and their frequencies of occurrences are not given.

10 The POS tags analysed can be viewed in Appendix 1; the complete UCREL CLAWS7 tagset can be accessed at: <http://ucrel.lancs.ac.uk/claws7/tags.html>.

analogous to Leggatt of *The Secret Sharer*. The double in the later work is Ransome [...]” ([6]: 53).

It may be argued that this finding might have been expected, in view of the stronger thematic and narrative similarities that TSS shares with TSL than HD. But what the analysis in Wmatrix provides is important quantitative validation of such a prediction, and (as we shall see with concordances) detailed, systematic insight into how those interconnections play out in the text. In spite of the features that the two texts share against which TSL is compared, the results in Table 1 show that statistically significant differences occur in the use not only of first-person pronouns (PPIO1: 1st person sing. objective personal pronoun [*me*]; PPIS1: 1st person sing. subjective personal pronoun [*I*]), but also third-person pronouns (PPHS1: 3rd person sing. subjective personal pronoun [*he, she*]). Similarly, another key pronominal POS-tag here, APPGE, (possessive pronoun, pre-nominal) represents pronouns with first person and third person, as well as second person, referents (e.g. *my, your, our, his/her*). If we unpack these POS tags, we can observe that *your, our* and *she* hold a much lower frequency of occurrence compared to the over-represented *I-he* words within the same POS tags. Considered that the story revolves around the masculine activity of a sea voyage, the pronouns such as *your, our* appear, in their co-text, to be mainly employed by the *I* narrator-protagonist to refer to a male or other males who, most of the time, is/are part of the crew. Instead, the *she* pronouns refer in most cases, as is the convention in English, to the ship the young captain left, and the ship he is appointed to. The over-represented narrative *I-he* pattern is further strengthened in the keyness analysis carried out at word level between TSL and the reference texts. The results are displayed in Table 2 in order of their statistical significance.

<i>TSL compared with HD</i>	<i>LL/+</i>	<i>TSL compared with TSS</i>	<i>LL/+</i>
my	42.54	me	19.47
he	28.83	he	12.54
me	24.44		
I	17.41		
his	16.59		

Table 2: The most-key pronouns in TSL compared with HD and TSS.

As already noted in the POS tag keyness analysis, higher statistical significance is recorded in the comparison between *The Shadow Line* and *Heart of Darkness*. The over-represented pronouns in the target text compared to the reference text (HD) show a significant use of *I* and *he* pronouns occurring in *The Shadow Line* and hint at a narrative centred around an internal self-referential *I*-voice with a strong focus on the external referential *he*. From the concordances produced by Wmatrix, the over-represented *I*-pronouns refer most of the time to the *I* protagonist, while the



*he* pronouns to the adults in his world,<sup>11</sup> especially to his ship-mates Mr. Burns and Ransome. This suggests that the young captain, during his difficult and challenging voyage, relies on a dyadic *I-he* type of interaction and tends to exclude *I-they* and the *I-we* relationships.

The dynamics of the *I-he* interaction can be better understood by observing their distribution across the narrative. For this purpose, in Section 5 I look at the concordance lines generated from three sample passages (cf. Appendix 2) taken from the beginning, middle and end of the story, respectively, each representing a key phase of the protagonist's maturing process.

### 5. Distribution of the over-represented *I-he* pattern in three samples from *The Shadow-Line*

Compared to the quantitative analysis of the pronouns carried out in Section 4, this Section is qualitatively-based as it makes use of the concordance tool in Wmatrix to search for the occurring *I-he* patterns in three key narrative moments of *The Shadow Line*. The concordance tool allows us to search for a word/node with a given amount of co-text to its left and to its right ([36];[1]). In order to look at the concordances for the pronouns in the sample passages, I uploaded each text-sample in Wmatrix, produced their concordance lines, and manually sorted out the lines in which the *I/he* pronouns occur as subject in order of occurrence in the samples. The only exception is found in Sample 1 where the occurrence of the first-person possessive ('my') is shown. Other pronouns occurring in the vicinity of the node (i.e. the search pronouns '*I*' and '*he*') in the concordance lines are also considered.

Text Sample 1 is taken from the opening of the story when the unnamed protagonist is telling the narratee (understood in this case, by default, as the reader) about the reasons for quitting his job as a mariner on a steamship. The passage exemplifies an element typically found at the beginning of a novel of formation in which, as noted by Stape and Simmons (in [14]: xlv), "the youthful hero [...] experiences a sudden discontent, shock, or impulse for change that entails a break with the known routines of family, career or traditional home." In the concordances produced, only *I*-pronouns appear since the *I*-narrator, in this segment of text, is the sole occupant of the narrative, which is characterised by an exclusive focalisation on the protagonist's inner self.

---

11 Special attention was paid to distinguishing between the *I*-pronouns referring to the *I*-protagonist and those referring to the other characters, occurring in their direct speech. For example, the pronoun *I*, though mainly indicating the narrator-protagonist, can also refer (approximately 83 times out of 1,301) to other characters that the *I*-protagonist interacts with in the story.

*Text Sample 1: I at the beginning of the story*

1. Only the young have such moments.	I	don't mean the very young. No. They
2. life in which such moments of which	I	have spoken are likely to come. What
3. story. It wasn't so bad as that with me.	My	action, rash as it was, had more
4. dissatisfaction. Rash moments.	I	mean moments when the still young are
5. a sensible person could put a finger	I	threw up my job, chucked my berth, left
6. put a gloss on what even at the time	I	myself half-suspected to be a caprice

After the general opening statement "Only the young have such moments," (l.1), the *I* presents himself as the speaker of the succeeding vague statement: "I don't mean the very young," (l.1). In ll.2 and 4 the *I* occupies a rather 'weak' position as it post-modifies "such moments of which *I* have spoken," "Rash moments. *I* mean moments." The speaker seems to hide himself behind intentionally generic observations. However, the speaker's leading role in the story is defined through the objective *me* (l.3), and the possessives *my*, and *myself* (ll. 5 and 6), which make the story more subjective and personal.

Text Sample 2 is taken from the middle of the story, when the young protagonist's trials reach their climax: the Chief Mate, Mr. Burns, is ill, the rest of the crew is also seriously ill, the ship is marooned in a dead sea, and thus unable to continue the planned voyage. Here, the young commander relies on his helper, Ransome.

*Text Sample 2: I-he in the middle of the story*

1. talk any nonsense to you of late? "	I	asked casually. "No, sir. "
2. this nonsense enough for you? "	I	asked, looking confidently at the <del>int</del>
3. by the agency of Captain Giles,	I	had yet an uneasy feeling that such
4. perhaps to be paid for in some way .	I	had held, professionally , a review
5. professionally, a review of my chances .	I	was competent enough for that.
6. competent enough for that . At least ,	I	thought so. I had a general sense of
7. for that . At least, I thought so.	I	had a general sense of my preparedness
8. the world . As natural as breathing .	I	imagined I could not have lived without
9. As natural as breathing . I imagined	I	could not have lived without it . I do
10. d I could not have lived without it .	I	don't know what I expected . Perhaps
11. lived without it. I don't know what	I	expected . Perhaps nothing else than
12. ce of youthful aspirations. Whatever	I	expected I did not expect to be beset
13. ful aspirations. Whatever I expected	I	did not expect to be beset by hurricanes
14. expect to be beset by hurricanes.	I	knew better than that . In the Gulf of
15. are no hurricanes. But neither did	I	expect to find myself bound hand and
16. passing his door. I believe that if	he	had had the strength he would have
17. that if he had had the strength	he	would have called out after me more
18. out after me more than once. But	he	hadn't the strength. Ransome, however
19. be picking up wonderfully. " " Did	he	talk any nonsense to you of late? "
20. question; but, after a pause,	he	added equably: " He told me this morning
21. He told me this morning, sir, that	he	was sorry he had to bury our late
22. is morning, sir, that he was sorry	he	had to bury our late captain right in
23. of care. Ransome didn't know.	He	had not given a thought to the matter
24. the matter. And with a faint smile	he	flitted away from me on his never-end
25. which only a man pursuing a calling	he	loves can know. That feeling seemed

The most frequently occurring pronouns are *I-he*; the pattern of occurrence in the concordances for *I-he* delineates a change in the narrator-protagonist's subject position within the narrative. The *I* is no longer focalising exclusively on his inner self but he is continually moving between a subjective internal position (*I*) and the external world by means of the subjective *he*, which deictically points either to Mr. Burns or Ransome, the antagonist and the helper respectively. From the concordance lines above, the *I*-protagonist (*ll*1-15) returns to play a dominant role with focalisation on himself, and he does not leave room, except in a very circumscribed way, for the external world (*ll*16-25) where focalisation shifts onto his preoccupation with the ill Mr. Burns (e.g. *ll*1, *ll* 16-19). Real or potential events (e.g. *ll*10, "I don't know what I expected") give rise to his reflections and impressions (e.g. *ll*13, "Whatever I expected I did not expect to be beset by hurricanes").

Text Sample 3 is taken from the final part of the story when, at last, the ship reaches the Port of Siam and the tension that has accumulated during the voyage is finally released: the crew is safe. Ransome unexpectedly informs the young captain of his decision to leave the ship, upsetting him. The ending of the ordeal foregrounds again the exclusive and somehow dependent relationship of the protagonist towards Ransome, as emerged in the analysis of Text Sample 2.

*Text Sample 3: I-he at the end of the story*

1. You don't mean to leave the ship !	"I	cried out. " I do really, sir. I want
2. to leave the ship! " I cried out."	<b>I</b>	do really , sir . I want to go and be
3. " I cried out. " I do really , sir.	<b>I</b>	want to go and be quiet somewhere.
4. <u>pit</u> will do. " " But , Ransome,"	<b>I</b>	said. " I hate the idea of parting
5. do. " " But, Ransome, " I said."	<b>I</b>	hate the idea of parting with you ."
6. the idea of parting with you ."	<b>I</b>	must go, " he broke in . " I have a
7. " I must go, " he broke in."	<b>I</b>	have a right! " . . . He gasped and a
8. he was another being. And	<b>I</b>	saw under the worth and the comeliness
9. about himself. " Of course	<b>I</b>	shall pay you off if you wish it, "I
10. shall pay you off if you wish it, "	<b>I</b>	hastened to say. " Only I must ask you
11. it, "I hastened to say. " Only	<b>I</b>	must ask you to remain on board till
12. board till this afternoon .	<b>I</b>	can't leave Mr. Burns absolutely by
13. understood that very well. When	<b>I</b>	returned on deck everything was ready
14. tempering my character--though	<b>I</b>	did not know it. It was awful. They
15. reproach of the bitterest kind, till	<b>I</b>	felt a sort of revolt wake up in me.
16. let them drop me, sir! " While	<b>I</b>	kept on shouting to him in most soothing
17. smile for a fleeting moment .	<b>I</b>	left for the shore in the steam pinnacle
18. parting with you. " "I must go, "	he	broke in . "I have a right !" . . .
19. broke in. " I have a right!" . . .	He	gasped and a look of almost savage
20. over his face . For an instant	he	was another being . And I saw under
21. him--this precarious hard life, and	he	was thoroughly alarmed about himself
22. by himself in the ship for hours."	He	softened at once and assured me with
23. in his natural pleasant voice that	he	understood that very well. When I
24. Frenchy had gone suddenly under.	He	was carried past me insensible, his
25. swollen, breathing <u>stertorously</u> ,	He	looked more like Mr. Punch than ever
26. contrary, had improved temporarily.	He	insisted on walking on his own feet
27. on each side of him. But	he	gave way to a sudden panic at the mom
28. out his weirdness amazingly .	He	looked like a frightful and elaborate

The most frequently occurring pronoun *I* refers mainly to the *I*-protagonist except for ll. 2-3, ll. 6-7 (in bold), which refer to Ransome. Here, we can see how the *I*-protagonist's external use of *I* in his dialogue with Ransome, marked out by the inverted commas in the concordance lines and in his internal thought, are centred on the anxiety and fear of having to depart from *he*/Ransome, the adult who has supported and guided him into his entry in the adult world. Through the second most frequently occurring pronoun *he*, the *I*-protagonist shifts his focus onto the external world, *i.e.* all his crew-members (Gambriel the grizzled sailor, Franky, and Mr. Burns). Now, the narrator-protagonist seems more absorbed in the needs of the people he is responsible for than in himself and this shift in attention to the real world could be a sign of his growing maturity.

## 6. Conclusions

In this article I have carried out a computer-aided analysis of the personal pronouns in Conrad's *The Shadow Line* compared to two other works by the same author. The WMatrix software was employed to identify and display the pronouns in order of statistical keyness in the target text compared to the reference texts. I then looked at the most frequently occurring pronouns in three different key phases of the story through concordances generated by Wmatrix. This analytical step exemplified how, in this work of fiction, the foregrounded *I*-protagonist is positioned in the external world during his rite of passage.

Further investigation could be carried out in order to ascertain to what extent the over-represented self-referential *I* and the other-referential *he* is a stylistic feature of the *Bildungsroman*, or Conrad's fingerprint. Also, the study of the key pronouns in the *The Shadow Line* through the frequently repeated sequences of words, bundles/clusters ([9]:989;[25]) would further contribute to showing the narrator's interaction with the adult world and his psychological development in this novella of formation.

Though to practitioners of the advanced and sophisticated current corpus-aided methods working with keyness analysis, frequencies and concordances might be regarded as a traditional and superseded method, they are still a highly relevant support in stylistic analyses of literary texts. In this article, I have emphasised that the key POS-tag approach is an important extension of the widely-used keywords technique, but extended to a more abstract level of analysis (grammatical word class, rather than simply lexis). Given the evidence of my case study, key POS-tag analysis, when combined with qualitative analysis (e.g. concordances) is a method that holds promise for literary stylistics more broadly.

## References

- [1] Adolphs, S. 2006. *Introducing Electronic Text Analysis: A Practical Guide for Language and Literary Studies*. Abingdon: Routledge.
- [2] Anthony, L. 2014. AntConc [Computer Software] Tokyo: Wasca University.
- [3] Archer, D., Wilson, A. & Rayson, P. 2002. Introduction to the USAS category system, 1–37.
- [4] Balossi, G. 2015. A computer-aided approach to *I* and the World in Conrad's *The Shadow Line*. On-line Proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA) pp.1-18.
- [5] Balossi, G. 2014. *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's The Waves*. Amsterdam: John Benjamins.
- [6] Benson, C. 1954. Conrad's Two Stories of Initiation. *PMLA*, 69 (1), 46-56.  
DOI:10.2307/460126

- [7] Biber, D. 2011. Corpus linguistics and the study of literature: Back to the future? *Scientific Study of Literature* 1(1) pp.15-23. DOI: 10.1075/ssol.1.1.02bib
- [8] Biber, D., Conrad, S., Finegan, E., Leech, G. & Johansson, S. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- [9] Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: CUP.
- [10] Burrows, J. F. 1987. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Clarendon Press.
- [11] Conrad, J. [1917] 2006. *The Shadow-Line*. Electronic Version, [www.gutenberg.org](http://www.gutenberg.org); [EBook #451].
- [12] Conrad, J. [1910] 2009. *The Secret Sharer*. Electronic Version, [www.gutenberg.org](http://www.gutenberg.org); [EBook #220].
- [13] Conrad, J. [1902] 2009. *Heart of Darkness*. Electronic Version, [www.gutenberg.org](http://www.gutenberg.org); [EBook #219].
- [14] Conrad, J. [1917] 2013. *The Shadow-Line, A Confession*. The Cambridge Edition to the Works of Joseph Conrad. Stape, J. H., Simmons, A. S. (eds). Cambridge: CUP.
- [15] Damerau, F.J. 1975. The Use of Function Word Frequencies as Indicators of Style. *Computers and the Humanities* 9: 271-280.
- [16] Dunning, T. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1): 61-74.
- [17] Garside, R. 1987. The CLAWS word-tagging system. In *The Computational Analysis of English: A Corpus-based Approach*. Roger Garside, Geoffrey Leech & Geoffrey Sampson (eds), 30-41. London: Longman.
- [18] Garside, R. & Smith, N. 1997. A hybrid grammatical tagger: CLAWS4. In Garside, R., Leech, G., McEnery, A. (eds.). *Corpus Annotation: Linguistic Information from Computer Text Corpora*, 102-121. London: Longman.
- [19] Gibbons, A. & Macrae, A. (eds.) 2018. *Pronouns in Literature: Positions and Perspectives in Language*. Basingstoke: Palgrave MacMillan.
- [20] Hirsch, M. 1979. The Novel of Formation as Genre: Between Great Expectations and Lost Illusions. *Genre* 12 (3), pp. 293-311.
- [21] Hoover, D.L. 2017. The microanalysis of style variation. *Digital Scholarship in the Humanities* 32 (2): ii17–ii30.
- [22] Kestemont, M. 2014. Function Words in Authorship Attribution. From Black Magic to Theory? *Proceedings of the Third Computational Linguistics for Literature Workshop*, co-located with EACL 2014 - the 14th Conference of the European Chapter of the Association for Computational Linguistics (27 April 2014, Gothenburg, Sweden): 59-66.
- [23] Leech, G. 1992. Corpora and theories of linguistic performance. In Jan Svartvik (ed.), *Directions in corpus linguistics*, pp. 105-122. Berlin & New York: Mouton de Gruyter.

- [24] Leech, G. 2013. *Virginia Woolf meets Wmatrix*. *Études de stylistique anglaise*. URL: <http://journals.openedition.org/csa/1405> DOI: 10.4000/csa.1405.
- [25] Mahlberg, M. 2013. *Corpus Stylistics and Dickens's Fiction*. London: Routledge.
- [26] Mahlberg, M. & P. Stockwell. 2016. Point and CLiC: teaching literature with corpus stylistic tools. In McIntyre, D. & Archer, D. 2010. A corpus-based approach to mind-style. *Journal of Literary Semantics* 39(2):167-182.
- [27] Murphy, S. 2015. I will proclaim myself what I am: Corpus stylistics and the language of Shakespeare's soliloquies. *Language and Literature* 24 (4): 338-354.
- [28] Propp, V. 1968<sup>2</sup>. *The Morphology of the Folktale*. Trans. by Laurence Scott. Austin: University of Texas Press.
- [29] Rayson, P. 2008. From key words to key semantic domains. *International Journal of Corpus Linguistics*. 13 (4): 519-549. DOI: 10.1075/ijcl.13.4.06gray
- [30] Rayson, P. 2003. *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. Ph.D. thesis, Lancaster University.
- [31] Rayson, P. Berridge, D. & Francis, B. 2004. Extending the Cochran rule for the comparison of word frequencies between corpora. *JADT: 7es Journées internationales d'Analyse statistique des Données Textuelles* : 1-12.
- [32] Schreihman, S., Siemens, R. & Unsworth, J. (eds) 2016<sup>2</sup>. *A New Companion to Digital Humanities*. Oxford: Wiley-Blackwell.
- [33] Scott, M. 2019. *WordSmith Tools version 8*. Stroud: Lexical Analysis Software.
- [34] Scott, M. 1997. *WordSmith Tools version 2*, Oxford: Oxford University Press.
- [35] Scott, M. 1997. PC analysis of key words - and key key words. *System* 25 (2): 233-245. DOI:10.1016/S0346-251X(97)00011-0
- [36] Sinclair, J. McH. 1991. *Corpus, Concordances, Collocation*. Oxford: OUP
- [37] Stubbs, M. 2005. Conrad in the computer: Examples of quantitative stylistic methods. *Language and Literature* 14(1), pp. 5-24. DOI: 10.1177/0963947005048873
- [38] Tabata, T. 1995. Narrative Style and the Frequencies of Very Common Words: A Corpus-Based Approach to Dickens's First Person and Third Person Narratives. *English Corpus Studies* 2: 91-109.

### Appendix 1: Pronoun tags analysed

APPGE	possessive pronoun, pre-nominal ( <i>e.g.</i> my, your, our)
PPGE	nominal possessive personal pronoun ( <i>e.g.</i> mine, yours)
PPHO1	3rd person sing. objective personal pronoun (him, her)
PPHO2	3rd person plural objective personal pronoun (them)
PPHS1	3rd person sing. subjective personal pronoun (he, she)
PPHS2	3rd person plural subjective personal pronoun (they)
PPIO1	1st person sing. objective personal pronoun (me)
PPIO2	1st person plural objective personal pronoun (us)
PPIS1	1st person sing. subjective personal pronoun (I)
PPIS2	1st person plural subjective personal pronoun (we)
PPX2	plural reflexive personal pronoun ( <i>e.g.</i> yourselves, themselves)
PPY	2nd person personal pronoun (you)

### Appendix 2

#### *TSL Sample text 1*

Only the young have such moments. I don't mean the very young. No. The very young have, properly speaking, no moments. It is the privilege of early youth to live in advance of its days in all the beautiful continuity of hope which knows no pauses and no introspection.

One closes behind one the little gate of mere boyishness--and enters an enchanted garden. Its very shades glow with promise. Every turn of the path has its seduction. And it isn't because it is an undiscovered country. One knows well enough that all mankind had streamed that way. It is the charm of universal experience from which one expects an uncommon or personal sensation - a bit of one's own.

One goes on recognizing the landmarks of the predecessors, excited, amused, taking the hard luck and the good luck together--the kicks and the half-pence, as the saying is--the picturesque common lot that holds so many possibilities for the deserving or perhaps for the lucky. Yes. One goes on. And the time, too, goes on - till one perceives ahead a shadow-line warning one that the region of early youth, too, must be left behind.



This is the period of life in which such moments of which I have spoken are likely to come. What moments? Why, the moments of boredom, of weariness, of dissatisfaction. Rash moments. I mean moments when the still young are inclined to commit rash actions, such as getting married suddenly or else throwing up a job for no reason.

This is not a marriage story. It wasn't so bad as that with me. My action, rash as it was, had more the character of divorce - almost of desertion. For no reason on which a sensible person could put a finger I threw up my job - chucked my berth - left the ship of which the worst that could be said was that she was a steamship and therefore, perhaps, not entitled to that blind loyalty which... . However, it's no use trying to put a gloss on what even at the time I myself half suspected to be a caprice.

### *TSL Sample text 2*

I avoided giving Mr. Burns any opening for conversation for the next few days. I merely used to throw him a hasty, cheery word when passing his door. I believe that if he had had the strength he would have called out after me more than once. But he hadn't the strength. Ransome, however, observed to me one afternoon that the mate "seemed to be picking up wonderfully."

"Did he talk any nonsense to you of late?" I asked casually. "No, sir." Ransome was startled by the direct question; but, after a pause, he added equably: "He told me this morning, sir, that he was sorry he had to bury our late captain right in the ship's way, as one may say, out of the Gulf." "Isn't this nonsense enough for you?" I asked, looking confidently at the intelligent, quiet face on which the secret uneasiness in the man's breast had thrown a transparent veil of care. Ransome didn't know. He had not given a thought to the matter. And with a faint smile he flitted away from me on his never-ending duties, with his usual guarded activity.

Two more days passed. We had advanced a little way - a very little way - into the larger space of the Gulf of Siam. Seizing eagerly upon the elation of the first command thrown into my lap, by the agency of Captain Giles, I had yet an uneasy feeling that such luck as this has got perhaps to be paid for in some way. I had held, professionally, a review of my chances. I was competent enough for that. At least, I thought so. I had a general sense of my preparedness which only a man pursuing a calling he loves can know. That feeling seemed to me the most natural thing in the world. As natural as breathing. I imagined I could not have lived without it.

I don't know what I expected. Perhaps nothing else than that special intensity of existence which is the quintessence of youthful aspirations. Whatever I expected I did not expect to be beset by hurricanes. I knew better than that. In the Gulf of Siam there are no hurricanes. But neither did I expect to find myself bound hand and foot to the hopeless extent which was revealed to me as the days went on.

### *TSL Sample text 3*

"You don't mean to leave the ship!" I cried out. "I do really, sir. I want to go and be quiet somewhere. Anywhere. The hospital will do." "But, Ransome," I said. "I hate the idea of parting with you." I must go," he broke in, "I have the right." He gasped and a look of almost savage

determination passed over his face. For an instant he was another being. And I saw under the worth and the comeliness of the man the humble reality of things. Life was a boon to him - this precarious hard life, and he was thoroughly alarmed about himself.

“Of course I shall pay you off if you wish it,” I hastened to say. “Only I must ask you to remain on board till this afternoon. I can’t leave Mr. Burns absolutely by himself in the ship for hours.” He softened at once and assuaged me with a smile and in his natural pleasant voice that he understood that very well.

When I returned on deck everything was ready for the removal of the men. It was the last ordeal of that episode which had been maturing and tempering my character - though I did not know it.

It was awful. They passed under my eyes one after another - each of them an embodied reproach of the bitterest kind, till I felt a sort of revolt wake up in me. Poor Frenchy had gone suddenly under. He was carried past me insensible, his comic face horribly flushed and as if swollen, breathing stertorously. He looked more like Mr. Punch than ever; a disgracefully intoxicated Mr. Punch.

The austere Gamhril, on the contrary, had improved temporarily. He insisted on walking on his own feet to the rail - of course with assistance on each side of him. But he gave way to a sudden panic at the moment of being swung over the side and began to wail despairingly:

“Don’t let them drop me, sir. Don’t let them drop me, sir!” While I kept on shouting to him in most soothing accents: “All right, Gamhril. They won’t! They won’t!”

It was no doubt very ridiculous. The blue-jackets on our deck were grinning quietly, while even Ransome himself (much to the fore in lending a hand) had to enlarge his wistful smile for a fleeting moment.

I left for the shore in the steam pinnace, and on looking back beheld Mr. Burns actually standing up by the tafftail, still in his enormous woolly overcoat. The bright sunlight brought out his weirdness amazingly. He looked like a frightful and elaborate scarecrow set up on the poop of a death-stricken ship, set up to keep the seabirds from the corpses.