УНИВЕРЗИТЕТ У БЕОГРАДУ

ФИЛОЛОШКИ ФАКУЛТЕТ

Марија М. Милојковић

ЕКСПЛИКАЦИЈА И ПРИМЕНА ЛОУОВЕ КОРПУСНО-СТИЛИСТИЧКЕ КОНТЕКСТУАЛНО-ПРОЗОДИЈСКЕ ТЕОРИЈЕ НА МАТЕРИЈАЛИМА ЕНГЛЕСКОГ И РУСКОГ ЈЕЗИКА

Докторска дисертација

Београд, 2019

UNIVERSITY OF BELGRADE

FACULTY OF PHILOLOGY

Marija M. Milojković

EXPLICATING AND APPLYING LOUW'S CORPUS STYLISTIC CONTEXTUAL PROSODIC THEORY USING ENGLISH AND RUSSIAN TEXTS

Doctoral dissertation

Belgrade, 2019

БЕЛГРАДСКИЙ УНИВЕРСИТЕТ

ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Мария М. Милойкович

ИЗЛОЖЕНИЕ И ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ КОРПУСНО-СТИЛИСТИЧЕСКОЙ КОНТЕКСТУАЛЬНО-ПРОСОДИЧЕСКОЙ ТЕОРИИ ЛОУ НА МАТЕРИАЛЕ АНГЛИЙСКОГО И РУССКОГО ЯЗЫКОВ

Докторская диссертация

Белград, 2019

Подаци о ментору и члановима комисије

Ментор:

Проф. др Јелена Вујић, Катедра за англистику, Филолошки факултет, Универзитет у Београду

Чланови комисије:

Датум одбране:

Изјава захвалности

Захваљујем се свом ментору, проф. др Јелени Вујић, на подршци и помоћи у разним фазама мог научног рада.

Експликација и примена Лоуове корпусно-стилистичке контекстуално-прозодијске теорије на материјалима енглеског и руског језика

Резиме

Циљ ове докторске дисертације јесте да се проучи и докаже научна валидност корпусно-стилистичке контекстуално-прозодијске теорије, коју је предложио Бил Лоу (Louw 1993, 2000, 2010a, 2010). Истраживање треба да покаже да велики, репрезентативан и избалансиран референтни корпус датог језика садржи довољну количину емпиријских података да омогући знатно прецизније читање ауторског текста него што је то било могуће у аналогном периоду. Потребно је доказати да је Лоуова теорија – самосталан и логички непротивречан научни систем, и да има потпуно самосталан комплет истраживачких инструмената који може податке добијене коришћењем других научних метода да подвргне анализи и провери сопственим поузданим и објективним средствима (Кожин 2011: 210).

Ово истраживање треба да размотри Лоуову хипотезу да „текст чита текст“ и да је „колокација инструмент за проналажење значења“ (Louw 2010a) – односно, да референтни корпус на нивоу колокације и уз фертијанску свест о контекту ситуације омогућава много потпунију и објективнију семантичку анализу ауторовог текста него људска интуиција. Треба да се докажу почетне претпоставке овог аутора (Louw 1993) да одступање ауторовог текста од контекстуалне и колокацијске норме која се садржи у референтном корпусу указује или на реторички обрт или на „неискреност“ аутора, која у ширем смислу треба да се тумачи као непотпуно слагање самог аутора са својом изјавом. Гради се „скала неискрености“ дискурса као конкретне манифестације датог језика, од клишеа и баналности преко „ауторске искрености“ и „ауторске неискрености“ и све до лажи као ауторове намере. Предлаже се хипотеза да се случајеви намерног лагања и клишеа у реалности преклапају, и да „скала неискрености“ на тај начин чини затворен круг. Истраживање треба да дефинише које место реторички обрти и стилске фигуре заузимају на овој скали. Такође се, примера ради, формулишу корпусно-засноване дефиниције неколико познатих стилских фигура. На овај начин би требало доказати да контекстуално-прозодијска теорија може да пружи увид у ауторову намеру на основу корпусних података, наиме да ли је аутор употребио реторички обрт, да ли је имао за циљ да превари саговорника, и да ли је својом изјавом пренео више значења него што му је то била намера или него што је тога био свестан у тренутку комуникације. Ове поделе на крају би требало да пруже увид у то шта тачно можемо сматрати квалитетним, а шта неквалитетним (баналним

или манипулативним) дискурсом. Осим семантичке анализе дискурса, стилских фигура и ауторове намере, ова теорија се примењује на истраживање инспирисаности квалитетног дискурса (што је више одступања од норме, утолико је и инспирисаност већа). Посебна пажња се поклања анализи преведених текстова. Пошто употреба референтних корпуса пружа много потпунији увид у све импликације текстуалног значења, треба искористити ове податке и проверити, да ли је преведени текст пренео већину ових импликација и да ли садржи елементе значења који одсуствују у оригиналу. Ово води ка евалуацији постојећих превода, али и ка формулисању препорука за будуће преводиоце.

Пошто Лоу тврди да заснива контекстуално-прозодијску теорију на филозофији језика, конкретно на Раселовој дефиницији логичког језика и Витгенштајновом схватању логичке форме, у раду се објашњава веза између филозофског схватања језика и емпиризма референтног корпуса. Истраживање показује да референтни корпуси могу послужити као илустрација мишљења ових филозофа, и можда га могу унапредити с обзиром на обиље емпиријских података које садрже.

Што се тиче метода истраживања, контекстуално-прозодијска теорија проучава значења у ауторском тексту на нивоу колокација. Колокације могу бити лексичке (у овом случају полази се од дефиниција које су дали Синклер (Sinclair 1991: 170) и Ферт (Firth 1957: 181) и лексичко-граматичке. У првом случају истраживање се заснива на семантичким прозодијама (Louw 1993) методом „wildcarding“ и на проучавању контекста у корпусу методом коселекције две или више лексема. Онда се даје мишљење о степену одступања аутора од колокацијске и контекстуалне норме у корпусу. Уколико се ради о лексичко-граматичкој колокацији, сви лексички колокати се апстрахују методом „wildcarding“, да би се онда све семантичке импликације датог граматичког низа проучиле у референтном корпусу кроз скуп његових најфреквентнијих лексичких колоката, које Лоу зове квазипропозиционим варијаблама датог низа, односно његовим корпусним подтекстом. Корпусном подтексту се у раду посвећује посебна пажња, јер је то ново откриће контекстуално-прозодијске теорије и није се детаљније проучавало у литератури, осим у радовима чији су аутори Лоу и Милојковић. У раду се наводи да би се примери у којима подтекст најављује мотиве који ће се касније појавити у тексту могли сматрати доказом да су референтни корпуси поуздан извор података о значењу текста које садрже граматички низови и које није доступно без употребе корпуса.

У раду се проучавају енглески и руски књижевни и некњижевни текстови, укључујући поезију. Главни референтни корпуси су British National Corpus (BNC), Corpus of Contemporary American English (COCA), корпус Google Books, корпус новина *Times* за 1995. годину и Национальный корпус русского языка (НКРЯ).

На почетку се даје преглед литературе чији је аутор Бил Лоу, као и постојеће литературе која коментарише и критикује контекстуално-прозодијску теорију или њене елементе (нпр. њено схватање семантичких прозодија). Након тога у раду се објашњава, на материјалима на енглеском и руском језику, како дата теорија треба да се примењује да би се дошло до дубљих увида у семантику ауторског текста, како се одступање од корпусне норме разликује од реторичког обрта, и који су тачно механизми на нивоу колокације који су одговорни за реторичке ефекте у тексту, од клишеа преко реторичких обрта (који могу али и не морају да се идентификују као познате стилске фигуре), и све до намерне лажи, с импликацијама за форензичку лингвистику. Објашњава се улога културног контекста. На крају, разматрају се филозофска дела која Лоу користи да би образложио своју теорију.

Посебна пажња се посвећује анализи превода, с руског језика на енглески. Ту као пример служе преводи поезије, пошто је ово форма где је концентрација значења од посебне важности.

Што се тиче употребе теорије у наставној пракси, рад укључује истраживања спроведена са студентима Катедре за англистику Филолошког факултета Универзитета у Београду, с двојаким циљем: да провери каква је рецепција ове теорије код студената друге и треће године и које би препоруке требало следити за упознавање будућих стручњака с овом методом. Ово истраживање укључује и квантитативни, и квалитативни део, и показује да је остварен одређени успех у настави корпусне стилистике и конкретно Лоуве теорије и методологије.

Кључне речи: корпусна стилистика, контекстуално-прозодијска теорија, семантичка прозодија, корпусни подтекст, неискреност, Лоу

Научна област: филолошке науке (лингвистика)
Ужа научна област: корпусна стилистика
УДК број:

Explicating and applying Louw's corpus stylistic Contextual Prosodic Theory using English and Russian texts

Abstract

The goal of this dissertation is to confirm the scientific validity of Louw's corpus stylistic Contextual Prosodic Theory (Louw 1993, 2000, 2010a, 2010b). The dissertation aims to show that a large, balanced and representative corpus of a given language contains sufficient empirical data to allow much more precise reading of authorial texts than it was possible in the analogue period. It aims to prove that Louw's theory is an independent and coherent scientific system and that it has at its disposal an independent set of tools, which can use its own reliable and objective procedures to verify data obtained by other scientific methodologies. (Кожин 2011: 210).

In particular, this study aims to verify Louw's hypothesis that 'text reads text' and that 'collocation is instrumentation for meaning' (Louw 2010a), which means that reference corpora at the level of collocation and in the light of Firthian context of situation enable much more precise and objective semantic analysis of authorial text than human intuition. One of its goals is to prove Louw's initial assumption that a deviation from the contextual and collocational norm in an author's text, established in the reference corpus, points either to a rhetorical device or to 'insincerity' on the part of the author, the latter broadly understood as less than total agreement on the part of the author with his or her own statement. To this end, a scale of 'prosodic and subtextual clashes' is introduced, classifying all discourse as an empirical manifestation of a given language, from cliché to 'authorial sincerity' and 'authorial insincerity' to intentional lying. A hypothesis is offered that intentional lying may turn into a cliché, thus transforming the cline into a circle. The research defines the place of rhetorical figures on the cline, and, on the basis of a few examples, offers corpus-assisted definitions of several known rhetorical figures. This is intended to prove that Contextual Prosodic Theory may shed light on authorial intention, i. e. on whether the author intended a collocation to act as a rhetorical device, to deceive the reader/listener, or whether it conveys a deeper meaning than the author may have intended or than he or she may have been aware of. These divisions are also intended to show what exactly may be considered quality discourse, and what may be empirically described as non-quality (banal or manipulative) discourse. Apart from semantic analysis of discourse, defining rhetorical figures and clarifying authorial intention, Contextual Prosodic Theory is applied to the study of inspired quality discourse, with the assumption that the more deviations from the norm are found in such discourse, the more inspired it is. Also, a special study on the analysis of translated texts will be included. As reference corpora offer

much deeper insight into textual meaning, the methodology of CPT should be used to check if the translated text contains the majority of the implications conveyed by the original, and if any of its meanings are not present in the original. This forms the basis of evaluation of existing translations, but also of a list of procedures recommended for use by future translators.

Louw's claim that he has based Contextual Prosodic Theory on philosophy of language, in particular on Russell's definition of logical language and Wittgenstein's understanding of logical form, makes it necessary to explain the connection between philosophical understanding of language and the empiricism of the reference corpus. It is shown that reference corpora may serve as an illustration of these philosophers' thinking, and arguably to further it through the abundant empiricism they offer.

As for methodology, CPT studies meanings in authorial text at the level of collocation, whether lexical or lexico-grammatical. As for the former, collocation is understood in the sense it was defined by Sinclair (1991: 170) and Firth (1957: 181). Its meanings are studied through present and absent collocates (established through semantic prosody) by means of wildcarding and co-selection in the context of situation. This is how the degree of the deviation from the corpus norm in authorial text is established. As for lexico-grammatical collocation, all lexical collocates of grammar strings are wildcarded so that its semantic implications may be studied in the reference corpus as the list of its most frequent lexical collocates, termed by Louw its quasi-propositional variables, or its corpus-derived subtext. Corpus-derived subtext is treated as an area of special interest, since it is a new development in CPT and has not been studied at length by authors other than Louw and Milojkovic. In particular, the ability of subtext to prospect motifs that are to follow in authorial text could be viewed as proof that reference corpora may be consulted in order to discover meanings in texts that are carried by grammar strings and are inaccessible to human intuition.

This dissertation takes into account English and Russian literary and nonliterary texts, including poetry. The main reference corpora used are the BNC, COCA, Google Books corpora, the corpus of the 1995 edition of the Times newspaper and Russian National Corpus.

After an initial overview of the works published by Louw and of the existing literature which either comments or criticises CPT or some of its notions (e.g. its understanding of semantic prosody), the dissertation focuses on illustrating, using English and Russian material, how CPT may be used in order to shed light on the semantics of authorial text, how a mere deviation from the reference corpus norm differs from a rhetorical device, what particular mechanisms at the level of collocation are responsible for rhetorical effects in a

text, from clichés and turns of phrase (these may or may not be identified as known rhetorical figures) to intentional lies, with implications for forensic linguistics. In the end, an account of philosopher's works upon which Louw founds CPT will be given, illustrating in detail the connection between philosophy and corpus stylistics. When it comes to the analysis of Russian-English translation, it will be based on poetic texts, as it is a genre in which concentration of meaning is of the utmost importance.

As for the use of CPT in teaching practice, the dissertation also contains an account of research conducted with students of the English Department at the Faculty of Philology, University of Belgrade. The aim of this study is two-fold: to check how CPT may be received by second and third year students of English, and what sort of methodology ensures most success in a corpus stylistic classroom in order to educate future experts. The study shows that some considerable success has been achieved on this first occasion of introducing CPT to university students.

Key words: corpus stylistics, contextual prosodic theory, semantic prosody, corpus-derived subtext, insincerity, Louw

Field of study: linguistics
Subfield: corpus stylistics
UDC:

Contents

1. Introduction…….....…...1
1.1 The academic context in which Contextual Prosodic Theory emerged…….....…..1
1.2 The main premises of Contextual Prosodic Theory…….....6
1.3 The aims and the methodology of the study…….…..….....11
1.4 The structure of the thesis……...….....12
1.5. Conclusion……..…….....…16
2. The work of Bill Louw to date…….....…17
2.1 Principles of corpus stylistics as established by Louw…….....17
2.2 Early work by Louw: relexicalisation and semantic prosody…….…….....….17
2.2.1 Delexicalisation in the study of literary texts…….....17
2.2.2 Irony or insincerity as the two main outcomes of breaches of semantic prosody…….....18
2.2.3 Outlines of Contextual Prosodic Theory defined…….....…20
2.2.3.1 Semantic prosody or connotation?.....20
2.2.3.2 The role of Firthian context of situation in establishing Contextual Prosodic Theory…21
2.2.3.3 Defining Contextual Prosodic Theory……...…….22
2.3 Forensic applications: exploring lexico-grammatical sequences in manipulative discourse and the language of negotiation…….....22
2.3.1 'Dressing up waiver': repeatable (and otherwise) language events in the reference corpus…….....…….....…22
2.3.1.1 Collocation as instrumentation for meaning and a dictionary of events….....23
2.3.1.2 The reference corpus as a source of 'events'…….....…24
2.3.2 The language of negotiation in specialised corpora…….....…24
2.4 Literary worlds as collocation…….....…25
2.4.1 Jacobson's poetic function replaced by the 'focusing' function of collocation.....25
2.4.2 The scientific nature of collocation……...…..26
2.4.3 Co-selection in establishing states of affairs…….....26
2.5 Towards corpus-attested definitions of literary devices…….....27
2.6 Corpus-derived subtext…….....….28
2.6.1 The beginnings of subtext: the logical construction of the world……...…..28
2.6.2 Corpus-derived subtext as 'instrumentation for meaning'……...……..29

2.7 The justification of inductive reasoning in corpus stylistics……..30
2.8 Corpus-derived subtext as prospection……..31
2.9 Proximity as entailment…….….33
2.10 Conclusion.....33
3 Semantic prosody in academic discourse, with particular emphasis on the work of Louw…36
3.1. The work of John McHardie Sinclair……36
3.1.1 Introduction…….36
3.1.2 The basic premises of Sinclair's approach to meaning……..…….….36
3.1.3 Collocation, colligation, semantic preference, semantic prosody……..…..…40
3.1.4 Sinclair on prospection……..…41
3.1.5 The word vs. a unit of meaning……41
3.1.6 Conclusion……42
3.2. The perception of semantic prosody in corpus linguistics……..42
3.2.1 Refinement and re-naming of semantic prosody……..44
3.2.2 Partington (2004)……..45
3.2.3 Whitsitt (2005)……..46
3.2.4 Hunston (2007)……..47
3.2.5 Morley and Partington (2009)……...……48
3.2.6 Stewart (2010)……...……48
3.2.7 McEnery and Hardie (2012)……...……51
3.2.8 McIntyre (2018)…….……53
3.2.9 Toolan's views on semantic prosody and his treatment of prospection…….....….55
3.2.9.1 Toolan (2007)…….....…55
3.2.9.2 Toolan's work on prospection…….……56
3.2.10 Philip (2011)……58
3.2.11 Concluding remarks…….58
3.3. Louw's impact on the discipline of corpus stylistics……60
3.3.1 Stylistics and meaning……60
3.3.2 What is corpus stylistics?.....62
3.3.3 Summary…….....…65
4. How ‘shareable’ is textual meaning?.....67
4.1 Is meaning in the text or in the reader?.....67
4.2 Readers' responses to foregrounding: experimental studies…….....….68

4.3 Are meanings dependent on patterns: a linguistic perspective……...….69
4.4 Research of semantic prosody in experimental psychology……...…….70
4.5 The interaction of semantic auras within a text……...……72
4.6 General semantics or psychological association?.....77
5 Transferability of Contextual Prosodic Theory to Russian……....……79
5.1 Introduction……...…….79
5.1.1 Semantic prosodies in Russian……...……..79
5.1.2 Semantic prosodies in Serbian……...……83
5.1.3 States of affairs, events, transitions……...…..85
5.1.4 Relexicalisation and literary devices…….……...….87
5.2 A prosodic clash in Pushkin.....89
5.2.1 The analysis of *ожившил войной*……...….89
5.2.2 Authorial intention in *ожившил войной*……...…….92
5.2.3 Relexicalisation, authorial attitude, and device……...……..……93
5.3 Delexical subtext vs. subtext relexicalised: which one is part of a device?.....94
5.4 Establishing subtext in Russian……....…..101
5.4.1 The author's collocation and the evidence……...…....101
5.4.2 Relexicalisation, co-selection and states of affairs in establishing subtext.....…106
5.4.3 Researching a shorter grammatical string for quasi-propositional variables...…115
5.4.4 The implications of the most frequent QPVs in **ась перед*……....…122
5.4.5 Pushkin's use of the verb *являлась*……...…..128
5.4.6 Establishing the subtext of *передо мной явилась ты* ('you appeared before me')…130
5.5 Literary device and authorial intention in Pushkin's *передо мной явилась ты* ('you appeared before me')…….…133
5.6 Implications for existing scholarship on authorial intention…….…..135
5.7 Conclusion…….…..136
6 Automating translation……138
6.1 Introduction…….138
6.2 Analysing an existing translation.....138
6.2.1 Subtext in the existing translation……...……..139
6.2.2 States of affairs in the existing translation……...…….158
6.3 Creating a corpus-based translation…….……162
6.3.1 States of affairs created by co-selection…….....…….164
6.3.2 Using parallel corpora when producing corpus-based translation….....…….167

6.3.3 Deciding on subtext in corpus-based translation……..…169
6.3.4 Corpus-based translation versus an existing professional translation.....172
6.3.5 Concluding remarks on sections 6.2 and 6.3.....177
6.4 Examining Falen's translation of *передо мной явилась ты*…….....180
6.5 Future directions…….....185
7 The interaction of semantic auras in a poetic text…….....187
7.1 Introduction…….....187
7.2. 'The Circus Animals' Desertion'.....187
7.2.1 The prelude: ‘I sought a theme and sought for it in vain’……...189
7.2.2 The middle stanzas: ‘What can I but enumerate old themes’.....193
7.2.3 The coda.....197
7.3 The subtext in the key metaphor of 'The Circus Animals Desertion'…….....204
7.4 Pun……..210
7.5 Personification in a poem by Brodsky…….…...216
7.6 Conclusion……..…221
8. Forensic linguistics: diagnosing manipulative discourse…….…….223
8.1 Introduction……...….223
8.2 Aims and methodology of the study……...…...224
8.3 Results of the study……..…..225
8.3.1 The context of the study…….....…225
8.3.2 The analysis of excerpt 1…….....…229
8.3.3 The analysis of excerpt 2…….…....233
8.4 Conclusion…….…..…242
9 Inspiration and authorial (in)sincerity.....244
9.1 Introduction.....244
9.2 The cline of prosodic/subtextual clashes within contexts of situation……..….246
9.3 Banality, sincerity, inspired sincerity in David Lodge's *Thinks* and *Changing Places*…247
9.4 Inspired insincerity in *Changing Places*.....277
9.5 Authorial insincerity, forced expression and manipulative language: quality vs. non-quality discourse……...……289
9.6 The place of rhetorical devices on the ‘insincerity cline’…….....…...291
9.7 Conclusion…….....…...….292
10 Inspired writing: a case in point…….....…...…..294
10.1Introduction…….....……294

10.2 A case study: inspired writing in English…….....……295
10.3 Conclusion…….....…….312
11 Contextual Prosodic Theory in the stylistics classroom…….....……315
11.1 Introduction…….....…….…315
11.2 The first classroom stylistic study…….....……..316
11.2.1 Aims……..……..…316
11.2.2 Methodology……...…….…..317
11.2.3 Background and limitations of the study…….……..….319
11.2.4 Implications of CPT for lexicography and classroom stylistics……...….....321
11.2.5 Quantitative research……..…323
11.2.5.1 The learning phase……..….324
11.2.5.2 The testing phase…….…....330
11.2.5.3 Discussion…….…..334
11.2.6 Qualitative research……...…337
11.3 Conclusion…….…...339
12 Student-centred stylistics: does subtext read text?.....343
12.1 Introduction…….……..…343
12.2 Aims…….…..…344
12. 3 Methodology…….…..…..345
12.4 Background to the study and its limitations…….……346
12.5 The testing phase……..….348
12.5.1 Stanza 1…….…….348
12.5.2 Stanza 2…….…….352
12.5.3 Stanza 5.....357
12.6 Relexicalisation as part of the testing phase…….....…..…..360
12.7 Conclusion to the testing phase…….....…...361
12.8 Students' feedback…….....…….…362
12.9 Conclusion…….....……...364
13. Prospection in novel-writing…….....…...….368
13.1 Introduction: research of prospection in short poems…….....…...…368
13.2 Prospection in students' essays…….....…..…369
13.3 Aims and significance of the study…….....…..….374
13.4 Lexical features of the texts under investigation…….....…..…376
13.5 Corpus-derived subtext in Faulkner's excerpt…….....…..…379

13.6 Corpus-derived subtext in DeLillo's excerpt…….....….389
13.7 The comparison of Faulkner's and DeLillo's opening passages….....…….396
13.8 Implications for literary criticism…….....……397
13.9. Conclusion…….....….401
14 Corpus-derived subtext and the context of culture…….....…403
14.1 Introduction…….....….403
14.2 Context of culture as a tool in a corpus stylistic analysis…….....…404
14.3 A corpus-attested view a context of culture…….....….408
14.3.1 The context of culture and the grammar of its lexis.....408
14.3.2 Method.....409
14.3.3 The context of culture in Faulkner's contexts of 'south' and 'southern'.….412
14.3.4 Conclusion…….....413
15 Is the truthfulness of a proposition verifiable through access to reference corpora?...415
15.1 Introduction.....415
15.2 The role of reference corpora in understanding and verifying philosophical concepts.....415
15.3 Wittgenstein: On Certainty.....417
15.4 Wittgenstein and Brodsky, or How much Brodsky knew about language games....421
15.5 Shared logical form……..……..425
15.6 Conclusion: prospection as a form of insincerity…….....……..431
16 Discussion ……..…..….434
17 Conclusion…….....…445
18 References…….....…..449
19. Биографија аутора….....465

1. Introduction

1.1 The academic context in which Contextual Prosodic Theory emerged

Corpora are 'large and principled' (Biber 2011: 15) collections of texts available to the researcher in a machine-readable form. Whether the researcher uses corpora available online (these are usually general corpora of a certain language, with specialised sub-corpora of that language), or a corpus they have compiled for their own research purposes, the corpus they are using comes with an interface, or a software package, that enables the researcher to study all instantiations of a certain language pattern in the corpus. The advantage of this method is that analogous research conducted manually is either impossible or could take an inordinate amount of time. The computer can quickly perform various statistical operations with the instantiations gathered (e.g. how frequent they are in the language, in what sort of surroundings they normally appear), which leads to insights concerning the typical behaviour of these patterns in context. Corpus linguistics, broadly defined, is a discipline which uses corpora in the study of language as its main source of data, and which is concerned with the building and development of corpora. To serve its purpose, a corpus needs to be representative – it can be 'evaluated for the extent to which it represents "a population" (ibid.).

Apart from these broad premises, the discipline of corpus linguistics has had a history of debates as to on what assumptions corpora should be built, and how and for what purposes they are to be consulted. Chomsky, for example, conveyed to John Sinclair, one of the leading figures in British corpus linguistics in the 20th century, that he dismissed the idea of corpora as such and that he carried a corpus in his head (Louw, personal communication). This stance has been documented (McEnery and Hardie 2012: 25) as Chomsky's choice to rely on experiments and not on vast collections of data, using which he appeared to consider impracticable. Despite Chomsky's reservations, nowadays corpus linguistics not only flourishes as a discipline, but has informed research in other disciplines, such as contrastive linguistics (Johansson 2007), discourse analysis (Baker 2006), language learning (Ajmer 2009), and sociolinguistics (Gabrielatos et al. 2010).[1] Indeed, saying that

[1] These examples are quoted in McEnery and Hardie (2012) as representative of research done in these areas. The present summary of the developments in corpus linguistics as an area is to some extent based on

corpora have become one of the major means of empirical verification of research findings in linguistics would not overstate the case. Nor would it be possible to draw a strict line between corpus linguistics as a discipline and other areas of linguistic study where corpora are used but whose practitioners do not consider themselves mainly corpus linguists.

The first collection of data, on file cards at first, was started at University College London in 1959 as part of the Survey of English Usage project (McEnery and Hardie 2012: 74). Inspired by the Brown corpus of written American English (Kučera and Francis 1967), this collection later evolved into the London-Lund Corpus of spoken language (Svartvik 1990), and LOB, FLOB and Frown, known as 'the Brown Family' (McEnery and Hardie 2012: 98-99). The International Corpus of English, known as ICE (Greenbaum 1996), comprising samples of a 'very wide variety of Englishes from around the world' (ibid.: 74) also originated at UCL. The UCL approach to corpus linguistics is characterised by a desire to offer a comprehensive grammatical analysis of the language, culminating in the publication of corpus-based grammars (see Quirk et al. 1972 and Quirk et al. 1985).

At Lancaster University, Geoffrey Leech developed tools for the searching, editing and annotating of corpus data. This resulted in the creation of corpora of around 2.8 million words, part of speech tagged and parsed. The fully automated annotation system made it possible 'to collect a corpus, annotate it while you had your lunch, and then work on the annotated corpus in the afternoon' (McEnery and Hardie 2012: 77). In response to Sinclair's lack of faith in human annotation, the Lancaster team proved (Baker 1997) that the consistency and accuracy were high (McEnery and Hardie 2012: 78).

On the other hand, a distinct approach to corpus linguistics originated at the University of Birmingham in 1960s and 1970s. It reached its heyday in 1980s, when Sinclair entered into partnership with the publishing house Collins and founded the research centre COBUILD (Collins-Birmingham University International Lexical Database). Sinclair's views were mainly influenced by those of Firth (1957) and partly Halliday (Sinclair 2004: vii). Sinclair's explicit reliance on empirical data and not pre-existing theoretical assumptions or speculation, following Firth's 'the meaning of the word *night* is

---

McEnery and Hardie (2012) and Hardie and McEnery (2010), particularly when it comes to non-neo-Firthian scholarship.

its collocability with *dark*' (Firth 1957: 196) as an explication of Wittgensteinian 'meaning is use' (OC 61-62, see also Philip 2011: 3-4), has led to the creation of not only an independent school of corpus linguistics, but an independent school of corpus linguistic thought, sometimes referred to as neo-Firthian. The COBUILD project resulted in the first corpus-based dictionary (the Collins COBUILD Dictionary), a corpus-based grammar of English (Sinclair et al. 1990), and in the creation of the Bank of English, a monitor corpus, which was meant to keep a daily record of every new word that appeared in the English language.

If at UCL the original focus was on grammar, at UB the focus was on lexicography – therefore, corpus work was directed at the search for meaning. Convinced that meaning was only definable in context, Sinclair's corpus methods involved studying a word/expression as the centre (the node) of a concordance (all or randomly selected contexts of the node in the corpus). The concordance is sorted in order of frequency, i.e typicality (Sinclair 1991: 105-106). The main approach adopted by Sinclair and the Birmingham school generally combined the manual reading of concordances for typical patterns with studying frequency lists and statistics of the collocates (co-occurring words) of the node. Such an approach, focused on how lexical items were associated with certain patterns in the language, resulted in Pattern Grammar (Hunston and Fransis 1999). An attempt to explain the connection between typical patterns and psychological associations took the shape of the theory of Lexical Priming (Hoey 2005), which will be discussed in Section 3.2. Thus, the Birmingham school mainly focused on the lexical items and their patternings, with models of grammar based on the collocative powers of lexis.

As already stated, Sinclair was not against using calculations of statistical significance (Sinclair et al. 2004: 28); a method combining reading concordances with quantitative data is explained in Stubbs (1995, 2014). Stubbs (2001: 71) advocates caution in the use of statistical methods, so do Hunston (2002: 78-9) and Moon (2007). Sinclair's viewing the human analyst as the supreme adjudicator when it comes to generalizing on the findings, and attaching great importance to what McEnery and Hardie call 'impressionistic' and 'hand-and-eye' methods (2012: 125) has given his opponents cause to pronounce that 'collocation-via concordance', unlike 'collocation-via-statistics' is 'the linguist's intuitive

scanning of the lines that yields up notable examples and patterns, not an algorithm or recoverable procedure' (ibid.: 126). These views, and in particular the notion that such an analysis is not replicable, will be critically examined in Section 3.2.7.

One of the main disagreements between these two schools of corpus linguistics is their view on annotation, with Sinclair considering that it obscured linguistic investigations because pre-tagging meant transferring pre-computational assumptions to the corpus (Sinclair's views on research methodology will be explained in Section 3.1). From this a theoretical stance emerged which in the literature has been termed 'corpus-as-theory' as opposed to 'corpus-as-method'. It has to do with the distinction introduced by Tognini-Bonelli of the Birmingham school: corpus-based vs. corpus-driven linguistics (2001: 66). While 'corpus-based' linguistics uses the corpus to find support for or refutation of a hypothesis formulated beforehand within a preconceived theoretical framework, 'corpus-driven' linguistics has no preconceptions. This same distinction has also been termed 'top-down' and 'bottom-up' (Gries 2010: 328). Sinclair used to tell his COBUILD team: 'Forget what you have learnt' (Louw, personal communication). This latter view has been understood as approaching 'corpus data in an exploratory fashion, i.e. without rigorously formulated hypothesis' (Gilquin and Gries 2009: 10, quoted in McEnery and Hardie 2012: 151). In fact, this approach by Sinclair simply cautions against pre-formulated bias when dealing with empirical data. However, it does mean that neo-Firthians do not approach corpus data from the starting point of a set of theoretical postulates, e.g. cognitive or pragmatic linguistic theories.

The search for meaning in context as the primary orientation of the Birmingham school led to collocation being viewed as co-occurrence, rather than as a result of statistical calculations. This approach alone yielded such a multitude of unexpected discoveries about language that it grew into a theory: the corpus was viewed as the only adjudicator, defying speculation and nullifying its importance, on the one hand, and undermining the status of statistical approaches because of their de-contextualised nature. Sinclair came to see collocation-based phenomena, in particular semantic prosody, as central to the study of language. This grew into the view adopted by Sinclairans that the corpus is not a method but has a theoretical status (Tognini-Bonelli 2001:1).

This is not to say that the scholars of the Sinclairan school have been unanimous in supporting each view outlined above, or that the school has never experienced disagreements from within. The most significant scholars that are normally viewed as part of the Sinclairan school are Hoey, Hunston, Krishnamurthy, Louw, Stubbs, Teubert and Tognini-Bonelli. These scholars have all adopted their own unique stances on the question of language and meaning. Hoey (2005) believes that meaning is in the mind and therefore ultimately inaccessible. Teubert (2010) claims that it is in the discourse, but with ever-changing discourse it becomes elusive. Krishnamurthy (2000) researches collocation using statistical software. Of these scholars, Bill Louw appears to hold particularly original views on corpus linguistics. A close associate of Sinclair's for several decades, and a determined Firthian rather than a neo-Firthian (Louw, personal communication), he saw his work on semantic prosody, a central concept in Sinclair's views on collocation, become a target of criticism starting from around 2005. Criticism of his seminal work, 'Irony in the text or insincerity in the writer: the diagnostic potential of semantic prosodies' (1993) came both from the neo-Firthian (e.g. Hunston 2007) and the 'corpus-based' school, but particularly harsh disapproval was expressed by academics at the University of Lancaster (McEnery and Hardie 2012, Xiao 2013). At the same time, after Louw (1993, 2000), little of his published work was seriously taken into account (with the exception of Philip 2011) until the publication of Simpson's *Stylistics* (2014), which practically revived Louw's reputation as a scientist in the Popperian tradition. This was followed by other mentions and positive reviews in the field of stylistics.

Louw's stance is characterised by his unswerving Firthian orientation and his uncompromising adherence to the main premises of the Sinclairan school. At the outset, Louw was only prepared to study raw untagged text and rejected the use of decontextualised statistics in principle. Instead, he sought confirmation of his views in the work of analytic philosophers: Frege, Russell, and Wittgenstein. Other reasons for the controversies surrounding his studies of semantic prosody stemmed from the simple fact that his place of residence was the turbulent political climate of Zimbabwe, where fruitful and accurate research was and still is impossible. Therefore, his works to date may have been too few and lacking in persuasiveness because they were focused on theoretical

assumptions and underexplained observations, with very few examples (e.g. Louw believes that he has proven his stance on the irony/insincerity dichotomy on the example of 'recipe for' in Louw 2000).

The goal of this dissertation is to offer a thorough review of the works by Louw and to place his work in a broader academic context, giving a detailed overview of all points of criticism that are known to have arisen since the publication of his early works. No one has so far given a systematic account of Louw's work between 2000 and 2019. It is only after a systematic review and justification of Louw's basic assumptions that fulfilling the other, no less important, goal of the thesis will be possible: to verify the principles behind Louw's Contextual Prosodic Theory (CPT) by applying it to the areas envisaged by Louw as suitable for its application. These are the study of transitions in texts, rhetorical devices, insincerity, forensic linguistics, teaching stylistic analysis, and translation. All these applications of this recently emerged and expanding theory, including the last and not formally (in print) foreseen by Louw, stem from Louw's dedication to the empirical study of meaning through collocation, an inherently Sinclairan concern.

1.2 The main premises of Contextual Prosodic Theory

While Teubert (2005) views growing reference corpora as a ground that leaves much to interpretation, claiming that linguistics is not like the natural sciences whose remit is the search for 'truth' (see Milojkovic 2013a: 60), Louw views reference corpora a means of verification. His claim that 'collocation is instrumentation for meaning' (Louw 2010a) has its roots in the notion of semantic prosody first elaborated by him in 1993. The theory, by means of which Louw claims scientific status for collocation, has developed further since then. I will briefly outline its main principles and developments (the outline will be based on Milojkovic 2013a).

Two search procedures in the reference corpus are crucial to verification: co-selection and wildcarding. Co-selection involves calling up all contexts in the reference corpus that contain the given two words. 'Wildcarding' is the designation (by means of *) of a point in the syntagm where a variety of words might on different occasions occur. The reference corpus may then be asked to produce all words that are used in the language in

the empty slot designated as *. For instance, searching a reference corpus for all occurrences of *brook * objection* may well return phrases where the middle word is one of a small set: *no, neither, the, any*, with *no* the most frequent variable in the reference corpus.

Contextual prosodic theory grew out of the study of semantic prosody (Louw 1993). Semantic prosody is an aura of meaning with which a word or phrase is imbued by its collocates, which means that collocates frequently occurring in the vicinity of a word or phrase will have 'shaped' that word's semantic aura in the mind or expectations of the native speaker. The word 'prosody' is an analogy: co-occurring words colour one another with meaning just like sounds transfer their qualities to adjacent sounds in phonology, for example, the nasal quality of the vowels in 'amen'. The first discovered semantic prosodies were mainly positive or negative; then further specificities were referred to. Since semantic prosodies are based on the frequency of consistent collocates, they can only be computationally recovered and are not accessible to intuition, although the native speaker will usually agree with computational findings when informed of them. Sinclair (2003: 117) refers to semantic prosodies as 'hidden meaning'.

Louw (2000) embeds the notion of semantic prosody and collocational mismatch into Firth's context of situation (Firth 1957: 182), and launches Contextual Prosodic Theory. The prosody is the result of a 'fractured' context of situation, i.e. a word or phrase is used more 'negatively' or 'positively' than usually, because Firth's context of situation helps interpretation as either 'overprovided' (more positive than the usual pattern found in reference corpora) or 'underprovided' (less so).

The notion of language events being interpreted through the context of situation (and Malinowskian context of culture) is connected to two important ideas. One is that Firth's definition of collocation involved its being abstracted at the level of syntax (Firth 1957: 196). What is asserted is its independence of syntactic restrictions: for example, 'dark' may make an appearance anywhere in the vicinity of 'night', even in the next or previous sentence.

The other important idea is that semantic prosodies, as well as their fractures and degree of fracture, are verified using large reference corpora. The bigger the difference between habitual usage and the poet's specific one – the greater the fracture – the more

defamiliarised the poet's usage. The Russian Formalists' expression 'ostranenie' comes to mind as a deviation from the existing norm, where the poet's usage is the deviation and the reference corpora contain the objective norm. However, while Douthwaite (2000: 37) defines foregrounding as a set of techniques deployed by the writer to draw the reader's attention to a particular part of a text, the writer does not need to be (nor can s/he be) conscious of the specific semantic prosody that her expression invokes in the reader, or of how fractured the semantic prosody of her expression may be. The aim of Contextual Prosodic Theory, by comparison, is to be able to research and explain, by using reference corpora, not only conscious but also *any* unconscious deviation from the language norm (i.e. a representative and balanced reference corpus) that the writer has employed. Thus, a corpus stylistician (see Section 3.3 for a discussion on corpus stylistics as a discipline focused on extracting meaning from texts using corpus linguistic tools) may select two lexical words and see what states of affairs will come up in a large reference corpus, and compare these to the state of affairs in the studied text. But states of affairs are not only created by co-selected lexical words. Grammatical strings also create states of affairs, and, consequently, prosodies (auras of meaning): for example, 'and now a' in a certain type of context of situation predicts a positive development in one's career (Milojkovic 2011a).

A grammatical string may contain lexical items. This is an important step in the development of CPT, as lexical items in a grammatical string may be *wildcarded* in order to see what state of affairs underlies the given string. Wildcarding is another corpus stylistician's tool for comparing the *event* in the target text to *similar events* in a large enough reference corpus via states of affairs created in them. Louw (2010a, 2010b) founds the notion of corpus-derived *subtext*, arrived at through wildcarding, on the work of analytic philosophers Russell, Carnap and Wittgenstein. Wittgenstein visualizes logical atomism within logical variables: 'Every proposition has a content and a form. We get the picture of the pure form if we abstract it from the meaning of the single words, or symbols [...] That is to say, if we substitute variables for the constants of the proposition, the rules of syntax which applied to the constants must apply to the variables also' (Wittgenstein 1929: 163). Here Wittgenstein supports Bertrand Russell's view on the language of logic, on which Louw bases his other analytic tool, subtext: 'A language of that sort will be

completely analytic and will show at a glance the logical structure of the facts asserted or denied . . . It is a language that has only syntax and *no vocabulary whatsoever . . . if you add a vocabulary*, [it] would be a logically perfect language' (Russell 2007: 197, emphasis added). It is through Russell's ideas that Louw came to call subtext 'logical semantic prosody'. To summarise, subtext is a state of affairs (Wittgenstein's *Sachverhalt*), or a literary world, arrived at by searching in a reference corpus a grammatical string found in a literary work or any other authorial text and wildcarding all content words. The most frequent content words shown by the corpus instead of the wild cards Louw calls, following Russell, quasi-propositional variables (QPVs).

This is how Louw (2010a) claims scientific status for collocation. However, some corpus linguists – Teubert (2005), Hunston (2007), McEnery and Hardy (2012) – seem to consider such interpretation at least partly subjective. Teubert claims that textual interpretation is negotiable, McEnery and Hardy – that it is unfalsifiable, and Hunston – that semantic prosody cannot be used predictively. Such objections will be dealt with in Section 3.2.

What follows is a schematic summary of the mechanisms involved in CPT (Milojkovic 2013a). A language event in the studied text (on the left) is compared to similar language events in the reference corpus (on the right). The basis for the similarity is Firthian context of situation (described in accordance with Firth 1957: 182). The transitions in text around which the language event is centred are forms of collocation (see sections 1.1 and 3.1 for Sinclair's view on collocation): lexical (checked in the reference corpus for semantic auras by co-selection) and lexico-grammatical (checked in the reference corpus for QPVs by wildcarding). Literary devices, though they may not be opaque to intuition, are forms of deviation from the (corpus-attested) norm (Louw 2008), as well as collocational mismatches that are less intuitively detectable. In fact, it could be said that every transition in text, being detectable at some level, is a device.

CONTEXTUAL PROSODIC THEORY

(CONTEXT OF CULTURE)	(CONTEXT OF CULTURE)
CONTEXT OF SITUATION	CONTEXT OF SITUATION

persons, personalities
verbal action of the participants
non-verbal action
relevant objects
effect of verbal action

LANGUAGE EVENTS

TEXT

persons, personalities
verbal action of the participants
non-verbal action
relevant objects
effect of verbal action

LANGUAGE EVENTS

REFERENCE CORPORA

COLLOCATION CHUNKS STATES OF AFFAIRS:

lexical collocation (co-selection)
prosodies of grammatical strings (e.g. 'and now a')
subtext (frequent QPVs within a grammatical string) (wildcarding)

Figure 1.1. The mechanisms of Contextual Prosodic Theory

To summarise, the tools of co-selection and wildcarding assist in discovering:

- semantic prosodies used in the studied text (auras of meaning of both lexis and grammatical strings) and prosodic clashes (telling mismatches between the author's and corpus usage), if there are such
- the subtext of key grammatical strings

The corpus stylistician will be capable of analyzing *events* in the studied text and *similar events* in the reference corpus, arrived at through co-selection and wildcarding, using Firth's context of situation. These tools will show how *states of affairs* are *chunked* in each case.

There is no true and fixed meaning, claims Teubert (2005). Still, a detailed analysis of text against the background of similar events in the reference corpus will narrow down the potential scope of interpretation with the closest resemblance to scientific exactness.

1.3 The aims and the methodology of the study

As stated at the end of Section 1.1, the main aim of this dissertation is to verify empirically the principles behind Louw's Contextual Prosodic Theory. This implies presenting evidence in favour of Louw's statement that collocation is instrumentation for meaning. To this end, several research questions must be answered. Their list is given below.

Is it possible to use reference corpora as a language norm against which particular authorial texts can be interpreted? If the answer to this question is positive, then is it legitimate to resort to inductive reasoning in textual interpretation?

In connection with this, can semantic prosody be explained and defended as a mechanism underlying the predictive quality of language?

How can the extraction of corpus-derived subtext contribute to nuanced textual interpretation?

Does deviation from the reference corpus norm result either in a rhetorical device or in insincerity?

How can corpus-attested definitions of rhetorical devices be obtained?

What is the precise connection between Russel's and Wittgenstein's logical language and natural language use as studied by Contextual Prosodic Theory? How can this link be illustrated?

This dissertation will aim to present Contextual Prosodic Theory as a rounded and logically coherent theory, with an independent set of tools (methods and terminology), capable of textual interpretation without the help of another theory, e.g. from the field of pragmatics.

As previously stated, this dissertation will also aim to illustrate not only how CPT can be applied to textual interpretation, especially as regards insincerity and rhetorical devices, but also how it is applicable to translation and forensic linguistics. Also, two studies in it will be reported: how CPT was first presented to university students of English, and, in particular, how students of English first reacted to the notion of corpus-derived subtext.

In order to prove the theory's transferability to other languages, it will be applied to Russian as well as to English - not only in the domain of translation, but also when it comes to rhetorical figures and insincerity.

The reference corpora used in the research described in the thesis are those on the website https://www.english-corpora.org/bnc/ that were available between 2011 and 2016 (sections 5-7, 9-12): BNC, 100 million words (Davies 2004-), COCA, 520 (earlier 450) million words (Davies 2008-). Findings obtained using corpora available on the website http://googlebooks.byu.edu/ will also be presented: Google Books - UK (34 billion words), Google Books - US (155 billion words), and Google Books - Fiction, 89 billion words (Davies 2011-). The corpus of the 1995 edition of the Times newspaper, 44.5 million words, has also been used (referred to as the 1995 Times corpus). As for Russian, the Russian National Corpus (RNC) with its subcorpora (Newspaper Corpus, Parallel Corpus, Poetry Corpus, Pushkin Corpus) has been used, its main corpus comprising around 170 million words in 2011, and reaching 280 million words in 2018. Searches described in sections 8, 13, 14 and 15 were performed in 2017-2018. The searches carried out were based on co-selection and wildcarding, and the analysis involved studying similar events using the context of situation and culture, as explained in the previous section.

It is presupposed that the BNC is a balanced and representative corpus for the purposes of co-selection and establishing semantic auras in British English. COCA was resorted to when the size of the BNC proved insufficient, or when studying American English. When it came to searching the subtext of grammatical strings, Google Books corpora were often consulted, as the size of the BNC and the COCA did not yield enough material in the majority of cases.

1.4 The structure of the thesis

The thesis consists of 17 sections. Section 1 briefly introduces the work of Bill Louw and defines the goals, methodology, and structure of the thesis, its main goal being the validation of his Contextual Prosodic Theory.

Section 2 presents Louw's work in more detail through the analysis of his publications. Particular attention is paid to the definitions of terminology as they emerge in his work, and which will be used throughout the thesis.

Section 3 places Louw's work in its broader academic context. First it surveys the premises adopted by corpus linguist John Sinclair, who may be considered both Louw's predecessor and his contemporary. Then it focuses on the critique of Louw's views as expressed in his publications and of approaches related to his. The views of academics who are primarily corpus linguists and those who are primarily stylisticians are discussed. This section will attempt to defend Louw's views against existing criticism and offer explanations for what is seen as misunderstanding or misrepresentation of his stance.

In Section 4 the question is posed whether textual meaning in the way it is interpreted by CPT can be understood by different readers similarly enough to be considered shareable. It presents results of those linguistic and psychological experiments that encourage an optimistic approach to this issue. The section then focuses on corpus-derived subtext and its prospecting qualities in order to illustrate that collocation is an empirical phenomenon and the basis of meaning being shared between the speaker/writer and the recipient.

Section 5 is the first of several detailed practical studies deploying the principles of CPT. They are applied to poetry authored by the Russian celebrated poet Alexander Pushkin, in this way offering evidence for the theory's transferability. As this is the first extended example of application, relevant terms are once again clarified. First semantic prosody, and then corpus-derived subtext are studied in Pushkin's poetic texts, which results in revealing authorial intention when it comes to conveying meaning. The borderline between rhetorical devices and authorial insincerity is discussed.

Section 6 proceeds to illustrate how CPT can be applied to translation. It is desirable that semantic nuances should be transferred to the target text, even if they are only accessible though reference corpora - at least, the semantic auras of the translation should not contradict those of the original. This section opens with the study of an existing partially successful translation of a line in one of Pushkin's poems, then describes creating a

corpus-based one, and finally analyses a professional translation offered by a celebrated translator of Pushkin's poetry.

Section 7 provides evidence that in a well-organised poetic text by a celebrated poet semantic auras of lexico-grammatical collocations interact and thus produce a complex impression in the reader. From the study of one of the poems by W. B. Yeats, and in line with the goals of the thesis, a corpus-attested definition of metaphor emerges. It is then compared to a corpus-derived definition of pun. The section ends with a corpus-based analysis of personification in one of the poems by a celebrated Russian 20th century poet Iosif Brodsky.

While the previous sections delineated the mechanism of extracting a deeper textual meaning than can emerge after an intuitive analysis, and showed how meaning as extracted from a poetic text may differ from the meaning that the author may have intended to encode (which may be considered a kind of non-manipulative insincerity), section 8 poses the question of how deliberate manipulation may be diagnosed in authorial text. It analyses a Russian newspaper article whose manipulative intention is manifest without recourse to corpora, and which has been found manipulative by a committee of Russian forensic experts. Application of co-selection and wildcarding in the RNC results in empirically revealing deeper mechanisms that aim at construing a false picture of reality in the minds of the readers.

Section 9 looks at a range of excerpts taken from the novels by the British writer David Lodge with a view to constructing a 'cline of prosodic and subtextual clashes'. The hypothesis is stated that specific deviations from the language norm may be connected with emotional fluctuations that are indicative of the state of inspiration – a state of mind which is characteristic of moments when quality discourse, be it poetic on non-poetic, is created. Taking into account the findings of sections 7 and 8, the places of rhetorical devices and manipulative discourse on the cline are also designated. In Section 10, the idea of a connection between certain types of deviation from the corpus norm and inspiration is further explored using poetry by the British published poet Jonathan Boulting.

Sections 11 and 12 take CPT into the domain of teaching corpus stylistics to university students. The aim of two studies described is both to investigate the subjects' reaction to the premises of CPT and to develop an adequate teaching methodology.

Section 13 is devoted to what is considered in this thesis as the main proof that corpus-derived subtext must be used in textual interpretation: prospection in the first lines of texts. First, examples of prospection in poetry referred to by Louw (Louw 2013, Louw and Milojkovic 2015, Louw and Milojkovic 2016) are analysed. Then a study of corpus-derived subtext in essays written by third-year students of English at the University of Belgrade is briefly described. Finally, a comparison of two opening paragraphs of novels written by American writers (one by William Faulkner and one by Don DeLillo) is presented, with conclusions that point to prospection existing in the first lines of a novel that can reach all the way to the text's concluding pages.

Section 14 discusses an aspect of corpus stylistic analysis that has not hitherto been given prominence: the context of culture. It shows how the context of culture may be taken into account in explaining why a textual segment has a certain impact on the reader, and how corpus stylistic analysis can deepen the reader's understanding of the context of culture of a novel (this time, again by William Faulkner).

Section 15 brings the truthfulness of authorial text back into the focus of the dissertation, this time from the point of view of philosophers of language Frege, Russel, and Wittgensten, whose views Louw considers analogue predictions of his digital instrumentation for meaning. Practical examples, a new one from Brodsky and a studied one from DeLillo, will be used to explain the link between philosophers' views on grammar and lexis and their manifestations in natural language.

Section 16 will discuss the findings of the thesis, particularly from the point of view of its stated aims, and Section 17 will contain its conclusion. It is hoped that aspects of Contextual Prosodic Theory that have not been given sufficient treatment in Louw's publications will thus have been explained from all relevant theoretical and practical angles, and that the empirical evidence presented in the thesis justifies Louw's theoretical thinking described in Section 2, and perhaps, when it comes to studies of inspiration and insincerity, even moves it forward.

1.5. Conclusion

This section has shown how Contextual Prosodic Theory, developed by Bill Louw, grew out of his cooperation with John Sinclair, the originator of the Birmingham school of corpus linguistics. This school is discussed in the literature as the neo-Firthian school because of its connection with the British linguist J. R. Firth, and represents the bottom-up, 'corpus-as-theory' and 'collocation-via-concordance' approach in corpus linguistic studies. Louw's theory, however, grew out of his studying particular authorial usages against the background of reference corpora, and developed from a focus on lexical collocation to studying its grammatical subtext, founding the latter on the work of analytic philosophers. The thesis will explain the premises and tools of CPT and illustrate their application to literary criticism, translation, forensic linguistics, and teaching stylistics to university students. The ultimate aim of the thesis is to present CPT as an independent theory capable of extracting meaning from texts, justifying Louw's claim that 'collocation is instrumentation for meaning' (Louw 2010a).

2. The work of Bill Louw to date

2.1 Principles of corpus stylistics as established by Louw

Corpus stylistics was first initiated by Bill Louw as a discipline at St Hilda's College, Oxford, in 1987 (Louw and Milojkovic 2014: 260). He was a close associate of John Sinclair through the collaboration of the University of Zimbabwe with the University of Birmingham via the British Council, and twice as a result was a visiting scholar collaborating on the COBUILD project (his name is mentioned in the Acknowledgements to the 2nd edition of the Cobuild dictionary). Having at his disposal COBUILD corpora of 22 million words, Louw saw the potential of the reference corpus to provide a more nuanced understanding of an expression's meaning, which could assist in a better understanding of a particular text. But of greater significance was Louw's insight that the corpus could also retrieve the frequent but absent collocates of a word. These collocates, frequent in the reference corpus, but absent from the particular text under study, were thought by Louw to play a role in our understanding of the text because of their frequency. The principle of taking into account absent frequent collocates underlies all main developments in Contextual Prosodic Theory: the claims of intuitive opacity of corpus findings, delexicalisation and relexicalisation as a basis of figurative language, the search for states of affairs in the reference corpus, for adherence to or breaches of semantic prosody, for corpus-derived subtext and prospection, and the probabilistic aspect of corpus stylistics.

2.2 Early work by Louw: relexicalisation and semantic prosody

2.2.1 Delexicalisation in the study of literary texts

Interestingly, the phenomenon of ***delexicalisation*** (see Sinclair 1991: 113) was first brought up in print by Low in the context of ELT rather than the study of literature or even classroom stylistics. The starting point of 'Classroom concordancing of delexical forms and the case for integrating language and literature' (1991) is Sinclairan lexical syllabus (Sinclair and Renouf 1987). It had been discovered that many very frequent English words had more than one meaning to them and that the majority of meanings were delexical: not literal in the sense that the full meaning of a word had been completely lost, such as of the

word 'take' in the expressions 'take a bus' or 'take a look'. Louw wondered at this point whether literary texts exhibited enough delexicalisation as a linguistic feature to be of use to a language learner. Using concordances from the Birmingham University corpus, he proves that this is indeed the case: words like 'ripple' and 'resurrect' are mostly used delexically in the language, even though, according to Louw, a great number of examples are drawn from literary texts. That is why, explains Louw, it is not enough for McNeice to use only one of expressions such as 'disinterred' or 'resurrected' in his poem ‘Refugees’ in order to create the central image of a group of people risen from the dead: in order for the metaphor to be sustained, several references to resurrection are necessary, otherwise the word 'disinterred' in McNeice’s ‘these disinterred from Europe’ will remain delexical! Appearing within the span of four words to the left and four to the right, which is Louw's preferred span, the members of this image cluster are here to ensure that the full meaning of these expressions stays full. Louw suggests that this intensification through reiteration is the very essence of ‘literariness’ as opposed to everyday language, which is not ‘device-bound’ to such an extent. This is the first mention of what Sinclair (2004: 198) will later refer to as 'a continuum of delexicalisation, crediting Louw. Nevertheless, in his analyses of how delexicalised meanings become literal in context, Louw (1991) in fact (although he does not use the term) describes the mechanism of ***relexicalisation***: the restoration of a word's original, literal meaning as part of the writer's intentions. This term will be used in Louw (2000) and subsequent publications (see Section 2.5.1).

2.2.2 Irony or insincerity as the two main outcomes of breaches of semantic prosody

Although Louw was not the first to have recorded in print the observation that semantic prosody (SP), which he defines as ‘an aura of meaning with which a word or phrase is imbued by its collocates’ (Louw 1993: 157), existed as a phenomenon (see Section 3.1), he was the first to use semantic prosodies as a key parameter upon which hinges our ‘subliminal’, according to Louw (ibid.), understanding of authorial text at the level of collocation. His seminal chapter ‘Irony in the text or insincerity in the writer: the diagnostic potential of semantic prosodies’ (1993) was the first to consider and utilise not the fact that SP exists per se (for the first description of the phenomenon, see Sinclair 1991:

75), but those cases where the author goes against the semantic tendency that can be established in the corpus. Deviations of necessity must exist, otherwise (if there are no exceptions to the rule, or semantic aura) we are not dealing with SP, but with connotation, observable at the conscious level (see Section 2.2.3.1). Drawing attention to such exceptions, Louw (1993) made an important and categorical 'either/or' claim: if the language user is not a foreigner or a child (that is, with insufficient language experience), his or her collocational clash is indicative either of intentional irony, or of unintentional insincerity.

As to irony, Louw (1993: 38) quotes the now famous example taken from *Small World* by David Lodge (1984: 4):

> The modern conference resembles the pilgrimage of medieval Christendom in that it allows the participants to indulge themselves in all the pleasures and diversions of travel while appearing to be austerely *bent on* self-improvement (emphasis added).

The SP of 'bent on' is negative. Such examples, according to Louw (1993: 41), solve some of the troubles Grice (1975) had when trying to pin down why we do not doubt that the utterance 'He is a fine friend' is ironic in the absence of any indicator apart from the tone of voice.

In Louw's single example of a collocational mismatch caused by lack of sincerity, the director of the British Council is interviewed by the Zymbabwean Television, who uses the word 'symptomatic' when praising the University of Zimbabwe. Unwilling to dismiss it as a slip of the tongue, Louw suggests the following interpretation:

> The speaker tries skillfully to avoid making the statement on Zimbabwe's national television that the University is dependent on Britain for assistance in the area of teaching and research [...] The full hypothesis may now be stated. Where encoders intend their remarks to be interpreted ironically, they 'write the device' in the form of an exception to an established semantic prosody. Conversely, where an utterance runs contrary to an established semantic prosody and it is clear from its context that

it is not intended by the encoder to be interpreted ironically, we find that 'the device writes the encoder' (Louw 1993: 44).

2.2.3 Outlines of Contextual Prosodic Theory defined

The next seminal paper by Louw, 'Contextual prosodic theory: bringing semantic prosodies to life' (2000)[2] was the first to formulate the name and essence of the theory that is the subject of this dissertation. The aims of Louw (2000), as stated in the paper, were not only to examine SP more closely as a phenomenon, but also to justify the link between linguistic and situational contexts of semantic prosodies, following Firth's (1957: 182) understanding of the context of situation.

2.2.3.1 Semantic prosody or connotation?

What differentiates text from visual content is, says Louw, the richness of its implications, apparently to a large extent underpinned by the phenomenon of SP, which is to a nowadays critic what connotation used to be in the analogue period. There are, however, two important criteria that are a distinct characteristic of SP as opposed to connotation: an SP is computationally recoverable, and, as an aura of meaning, does not cover all contexts of the given expression in the reference corpus.

What requires more explanation is, in my view, Louw's claim that SPs are 'more strictly functional or attitudinal'. Simpler put, whereas the presence of connotation is sufficient to indicate the speaker's attitude, semantic prosodies, whether adhered to or broken, ***chunk*** that attitude through ***co-occurrence***. Connotation always works as an association existing in the mind of the recipient (Louw gives the example of 'urchin'). In contrast, 'days' evokes no such association in the mind, but together with 'are' it chunks the right semantic construct to evoke the negative attitude that Larkin needs. We are fully apprised of his attitude because we have looked up 'days are' in the reference corpus and discovered that they are usually 'gone', 'past' etc. (Louw 1993: 160-163). Semantic prosody

[2] Louw kindly agreed to send me the final versions of the number of his publications that I had not been able to access in print (Louw 2000, 2006, 2007a, 2007b, 2008, 2010a, 2010b). These final versions are accessible to the public on https://uz-ac.academia.edu/BillLouw. I have included their description because they mark significant stages in the development of Louw's thinking.

is more functional than connotation because it would not have come into being and received its textual realization as a combination of separate linguistic building blocks in context but for Larkin needing that very specific semantic aura to materialize. To sum up, ‘more strictly functional or attitudinal’ must mean that SP only works if a word, any word at all, has a particular collocate within the Sinclairan nine-word window, while connotation always works as an association existing in the mind of the recipient.

Even if we do intuitively feel the mismatch – Louw’s example is 'fanned the flames of optimism' taken from Moon (1998: 161, 256) – the concordances are there to explain the mechanism fully: 'fan the flames of' is normally followed by negative collocates. SPs are contained even in the smallest grammatical building blocks such as articles: using 'a' instead of 'the', for example, or the forms containing 'some*', is likely to convey more sinister an impression because the entities described by these forms are indefinite, according to Louw.

2.2.3.2 The role of Firthian context of situation in establishing Contextual Prosodic Theory

The discussion of 'fan the flames' of and the forms of 'some*' brings to the fore the situational aspect of SP, not insisted upon in so many words in Louw (1993). This is Firth's comment on ***context of situation***:

> My view was, and still is, that 'context of situation' is best used as a suitable schematic construct to apply to language events, and that it is a group of related categories at a different level from grammatical categories but rather of the same abstract nature. A context of situation for linguistic work brings into relation the following categories:
>
> A. The relevant features of participants: persons, personalities.
>
> (i) The verbal action of the participants.
>
> (ii) The non-verbal action of the participants.
>
> B. The relevant objects.

C. The effect of the verbal action (Firth 1957: 182).

For example, Firth's 'persons, personalities' explains the sinister aspect of 'some*' forms: they are unknown. Reflexive forms like 'get *self *ed' betray the SP of disappointment. The situational profile is provided by the corpus, as in 'get herself pregnant'.

2.2.3.3 Defining Contextual Prosodic Theory

Having offered a revised definition of semantic prosody ('a form of meaning which is established through the proximity of a consistent series of collocates, often characterisable as positive or negative, and whose primary function is the expression of the attitude of its speaker or writer towards some pragmatic situation', Louw continues to insist on its connection with the context of situation, thus introducing Contextual Prosodic Theory. The terms 'fractured context of situation' and 'overprovision/underprovision of context' refer to how a collocational mismatch works at the contextual level, creating the deviation of the situational context of the utterance from the reference corpus norm describable as the sum of a word's collocates in the reference corpus. In short, the methodology of Contextual Prosodic Theory (CPT) in the year 2000 was based on checking authorial text at the collocational level against the reference corpus norm, seeking to establish that 'norm' both collocationally and contextually, following the description given by Firth (1957: 182).

2.3 Forensic applications: exploring lexico-grammatical sequences in manipulative discourse and the language of negotiation

2.3.1 'Dressing up waiver': repeatable (and otherwise) language events in the reference corpus

The property of lexico-grammatical sequences used in running text to reveal unexpected but very particular SPs was exploited by Louw to full advantage in his 'Dressing up waiver: a stochastic collocational reading of ''The Truth and Reconciliation Commission'' (Louw 2003). This time Louw's objective is not close analysis of literary

innuendos but corpus-driven study of the manipulative mechanisms behind creating ideological climates. While this pursues the 'insincerity' thread of the seminal 1993 paper, this time the forensic aspect of corpus analysis is explored to the full: the creators of the Truth and Reconciliation Commission (TRC) stand accused by Louw of intentionally manipulating public opinion. The repeatable events in the corpus are the logical link between the context of situation (or event) within authorial text and the one surrounding it (this time, the political situation in South Africa).

2.3.1.1 Collocation as instrumentation for meaning and a dictionary of events

Louw claims that the use of reference corpora proposed by Sinclair has turned linguistics into a hard science, with collocation having become digital 'instrumentation for meaning' (Louw 2003: 5). He advocates abandoning analogue description altogether in favour of corpus analysis that he calls data-assisted reading or corpus-based instrumentation (during which procedure some generalizations, strictly founded on fact, must of necessity take place, see Tognini-Bonelli 2001: 85). Meaning, claims Louw, is carried not by individual words but by their combinations instantiated in discourse in a particular context. (This, of course, does not preclude the mind from having associations.) The new type of dictionary, according to Louw, should be a dictionary not of words, but of ***events*** (clearly referring here to Firth's 1957: 182 'language events').

Louw's example in the paper is the term 'natural justice' (this example will later be quoted in Simpson's influential *Stylistics*, see Simpson 2014: 101). The reference corpus additionally supplies the vital semantic information that, in actual discourse, this term is most often used, as shown by its collocates to the left as well as to right, not when justice has been upheld but when it has been denied. Moreover, 'the narrative apparently ceases to the right of the expression. There is no continuation of a chained event with the word 'and' (Louw 2003: 14). This, according to Louw, sheds light on what is the 'normal' attitude of judges: when natural justice is denied, it remains denied. 'Justice is not natural in the same way yoghurt is natural' (ibid.: 16). This information we have obtained by considering collocation within Firth's taxonomy for the context of situation. Collocates to the right of

the node disclose the future course of events and present a less clear situation than those to the left.

2.3.1.2 The reference corpus as a source of 'events'

Still, the focus of this paper is not cases such as 'natural justice', where a sequence turns out to have a particular, situational, precise and unexpected meaning. This paper is devoted to cases when a certain collocation, and therefore meaning, does **not** exist in a reliable reference corpus. These cases are in fact artificially designed situations in which an artificially designed collocation comes into being to serve the purposes of the manipulator. The key example in this paper is the collocation 'Truth and Reconciliation Commission.'[3] Did 'truth' and 'reconciliation' collocate prior to the existence of TRC? According to Louw, the first large COBUILD corpora available in 1985 did not contain such collocations.

Louw refers to his earlier research on the language of the South African TRC, such as 'mass justice'. 'Mass' has a negative SP (collocating with 'murder' and 'destruction', see Louw 2000). Louw's verdict is that the language of TRC is false because it does not have precedents in the reference corpus. This subject will be further pursued in Louw (2007b), quoting Russell, who confirms the importance of taking account of the world in linguistic investigations: 'The linguistic philosophy, which cares only about language, and not about the world, is like the boy who preferred the clock without the pendulum because, although it no longer told the time, it went more easily than before and at a more exhilarating pace' (Gellner 1959: 15).

2.3.2 The language of negotiation in specialised corpora

The subsequent paper, 'Data-assisted negotiating: will it produce a new class of negotiator or destroy the ideology of negotiating?' (Louw 2006, reprinted as Chapter 4 in Louw and Milojkovic 2016) continues to explore the forensic aspect of corpora. Louw advocates the use of specialized corpora of negotiations in order to discover the subtle

[3] Ostensibly, this body was set up in South Africa in order to assist victims of apartheid. However, according to information supplied by Louw, the actual goal was not to remedy past hurts and issue compensation, but to use the lack of legal background on the part of the victims and make them waive their rights to compensation and go without 'natural justice'. Reconciliation was supposed to amount to forgiving the criminals. Victims demanded proper trials instead.

meanings of certain frequent expressions and facilitate not only the understanding of the negotiation process, but also the training of a ‘new class of negotiator’.

A convincing example offered by Louw is the expression 'just like that', used by an advertising firm in Zimbabwe to launch new software. In the reference corpus, reports Louw, the majority of the occurrences of the phrase in fact refer to failures and interruptions instead of successful operations. When it comes to negotiating, several frequently used expressions seem to have much subtler meanings than one would intuitively assume. For instance, 'correct me if I am wrong, but' is normally used when the speaker has checked her facts carefully and knows she is not wrong. 'Happy to' also belongs to the class of such expressions: the speaker is in fact sad rather than happy. 'Sorry to interrupt' undergoes the same reversal of meaning.

2.4 Literary worlds as collocation

2.4.1 Jacobson's poetic function replaced by the 'focusing' function of collocation

In his next several papers Louw revisits the impact collocation may and should have on literary criticism. In doing so, he integrates Jacobson’s structural analysis, Firthian context of situation and Sinclairan nine-word window of collocative power. ‘Collocation as the determinant of verbal art’ (Louw 2007a) opens with Jacobson’s definition of the poetic function: ‘The poetic function projects the principle of equivalence from the axis of selection into the lexis of combination’ (Jacobson 1960: 358). Louw explains that, focusing exclusively on syntax, Jacobson’s definition is too mechanical and does not in fact explain *how* the poet comes by his poetic choices. For example, Ezra Pound’s fourteen-word poem ‘In a Station of the Metro’ is a case in point as it does not contain any verbs at all:

> The apparition of these faces in the crowd
> Petals on a wet, black bough

If we were to follow Jacobson, says Louw, we would intuitively try to explain the choice of 'apparition' in 'the * of'. We might consult a thesaurus and discover equivalents such as 'appearance', 'manifestation' or 'chimera', 'phantom', and then speculate intuitively how these could have been considered and abandoned.

Louw does not ban intuition as such; he proposes to dispense with it as a scientific procedure. Corpora, and not dictionaries, hold the key to meanings. 'An apparition', says Louw, is scary because the corpus says it is literal. 'The apparition', according to the corpus data provided by Louw, is delexical and points to the appearance of the Virgin, as its most frequent collocates in the Bank of English at the time of writing were 'the' (57), 'of' (48), 'virgin' (9), 'an' (9), 'mary' (5), 'blessed' (3), 'saw' (3). 'Particular forms are dominated within the corpus by collocates which the reader cannot retrieve, either in the text that is being read, or intuitively from past experience' (Louw 2007a: 160).

2.4.2 The scientific nature of collocation

Thus, collocation provides computational means for the investigation of meaning in poetic discourse. In 'Truth, literary worlds and devices as collocation' (2007b) Louw wonders at to what extent Sinclair's plea contained in the title of his book *Trust the Text* (2004) has remained unheeded: no significant theoretical changes have occurred as a result of introducing corpora into linguistic studies. Louw proceeds to discuss the implications of Russell's (1960: 12-15) idea that, when philosophy settles an issue, it hands it over to science, and whether collocation as instrumentation for meaning is entitled to a scientific status. These ideas are further taken up in Louw (2011), 'Philosophical and literary concerns in corpus linguistics'. According to Louw, literary critics have always demanded to have an empirical foundation for the view of style as deviation from the norm. The moment empirical procedure for establishing what was norm and what was deviation was provided, such demands ceased (Louw 2011: 186). Louw quotes Sinclair as saying that 'we should trust the text. We should be open to what it may tell us. We should not impose our ideas on it' (Sinclair 2004: 23). Louw also reveals that since his seminal paper (Louw 1993) the type of deviation alternative to irony, namely insincerity, has not been researched at all.

2.4.3 Co-selection in establishing states of affairs

'Literary worlds as collocation' (Louw 2007c) continues to advance the case of computational literary studies. Apart from keyword analyses in the corpus of an author's

work, and from the study of SPs as fractured contexts of situation, here Louw also advocates resorting to ***co-selection*** in order to ascertain, for example, a poem's mood. He takes the poem 'My own heart' by G. M. Hopkins and co-selects 'self' and 'poor' in the BNC to discover that in the reference corpus they appear together in contexts of poor self-image, which in a nutshell explains the predicament of the poem's persona, who torments himself instead of showing himself more kindness (Louw 2007c: 102-103). The result of co-selection in the corpus is what Louw refers to in his next paper as ***states of affairs***: consistent types of usage in a particular type of context of situation in the reference corpus that, when compared to usage in authorial text under investigation, help interpret the text's meaning. In Louw (2008; see Section 2.5.2 for a fuller account) states of affairs are given a philosophical explanation:

> 1. The world is all that is the case.
> 1.1 The world is the totality of facts, not of things.
> 1.2 The world divides into facts.
> 2. What is the case – a fact – is the existence of states of affairs (Wittgenstein 1922: 7).

The role of collocation at the macro-level, claims Louw, is to determine the symbolism of a work of art by taking into account collocates that are frequent in the corpus but absent from the text (Louw 2008).

2.5 Towards corpus-attested definitions of literary devices

'Consolidating empirical method in data-assisted stylistics: Towards a corpus-attested glossary of literary terms' (2008) claims that the mechanisms of delexicalisation and relexicalisation can define literary devices, because relexicalisation is what all rhetorical devices have in common. Moreover, as stated in the title, Louw advocates the creation of a replicable corpus-attested glossary of literary terms in place of the existing analogue ones.

It was already stated in Louw (1991: 154) that a surprising number of words in the English language are used delexically more frequently than they are used literally. In Louw (2007a) the author explicates the mechanism of ***relexicalisation*** – reviving the literal meaning of a delexical expression by placing it in the vicinity of collocates which in the reference corpus co-occur with its literal instantiation. In other words, totally delexical expressions may relexicalise if they appear in the proximity of collocates that make their meaning literal in the corpus, as is the case in the opening line of Henry Miller's *Tropic of Capricorn:* 'Once you have given up the ghost, everything follows with dead certainty, even in the midst of chaos' (Miller 1967: 9). 'To give up the ghost' and 'dead certainty' are delexical expressions, but they are briefly relexicalised, or made literal, by their proximity. which is then proved by a reference corpus.

Louw (2008) is the last paper which dwells exclusively on semantic auras of lexis. Louw quotes Sinclair (2006) on the SP of the string 'when she was'. The collocates are 'hair-raising': 'diagnosed', 'assaulted' etc. The one that looks neutral, 'approached', either suggests physical attacks or excellent professional opportunities. Louw (2011: 190-192) follows up on Sinclair's search of 'when she was' and attempts the search of 'when he was'. He finds, for example, that the most frequent collocates of 'when he was in' are 'charge' and 'power'. That is why Louw places Firthian and Malinovskian context of situation embedded in the above the level of collocation, and lexicography below (Louw 2008: 247; Louw and Milojkovic 2016: 44, 53).

2.6 Corpus-derived subtext

2.6.1 The beginnings of subtext: the logical construction of the world

In 'The analysis and creation of humour' (Louw 2009)[4] approaches the breakthrough of corpus-derived subtext as a philosophy-based phenomenon that is about to emerge in 2010. Here Louw introduces for the first time, in addition to Wittgensteinian states of affairs, Carnapian ***'logical construction of the world'***. Verbal humour is seen both as a contextual and a collocational phenomenon. When Mona Baker, as a member of Sinclair's

[4] The text of this keynote address was kindly sent me by the author. It was also published as Chapter 5 in Louw and Milojkovic (2016).

team, wrote to a linguist asking to ‘get her hands on his corpus’, she was not aware that 'corpus' etymologically relexicalised her wish to ‘get her hands on it’. This example leads us to corpus-derived subtext because 'get * *s on' in the reference corpus (Louw uses BNC) yields only one lexical collocate, and that is 'hands'. Carnapian ‘logical construction of the world’ in this analysis is that it determines ***probability***, or likelihood, that a context of situation, expressed linguistically through a set of collocates, will ***entail*** a particular, and not some other result (e.g. that if the police is observed on the streets, it is likelier to keep watch and enforce law and order rather than to beg). Of course, in linguistic terms the probability is determined by the use of a certain collocation in certain contexts that can be grouped according to a definable state of affairs.

2.6.2 Corpus-derived subtext as 'instrumentation for meaning'

Louw's ‘Collocation as instrumentation for meaning’ (2010a) takes CPT further because collocation is now viewed not as a lexical, but as a lexico-grammatical phenomenon, and this is the first mention of ***corpus-derived subtext***. He analyses the line ‘a heart/ Made bleak by sacrifice’ from Herbert Williams’s poem ‘Daughter of the House’, only to find that the search line 'a * made * by' yields 'possible' as its lexical collocate in the second slot, highlighting not sacrificial, but rather opportunistic ('made possible by') nature of the protagonist’s desires. Louw adds (personal communication) that such a reading was confirmed by the poet himself.

It is in Louw’s next paper, ‘Automating the extraction of literary worlds and their subtexts from the poetry of William Butler Yeats’ (2010b) that corpus-derived subtext is defended in more detail. In accordance with Russell’s definition of logical language (see Section 1.2), Louw removes all vocabulary from the search string, leaving grammar only (the way he does with ‘Daughter of the House’ in the previous paper). His example this time is the first line in Yeats’s ‘Sailing to Bysantium’: ‘This is no country for old men’. He then enters the search string 'this is no * for' into the BNC and Bank of English and comes up with two most frequent lexical items: 'reason' and 'excuse'. The implication is that the persona travels to Bysantium against his better judgement. Thus, Louw considers the most frequent lexical variables within or around a grammatical string to be that string’s ***logical***

semantic prosody – or, ***corpus-derived subtext***. He calls these most frequent lexical variables ***quasi-propositional (QPVs)***, because any proposition in the natural language is traditionally referred to in philosophy as a 'quasi-proposition' (Louw, personal communication).

Summarising Louw's views on subtext as described in Louw (2010a, 2010b), it may be concluded that corpus-derived subtext is not only 'an abstraction at the syntagmatic level' (Firth 1957: 196), but is carried across the whole poem as an additional, implied meaning, invisible without recourse to corpora. This is what Carnap refers to as 'remembrance of similarity' (Mautner 2005: 86, quoted in Louw 2010a). The implication is that we are aware subconsciously of the frequencies of usage, be it the frequent but absent collocates of lexis or grammar, and these collocates have a bearing upon our interpretation of discourse. The meaning is so subtle it is below the radars of our intuition, because the semantic auras of grammar strings are completely opaque to intuition, while when it comes to lexis we can still talk of 'twenty-twenty hindsight' (Louw 1993: 173; see also Sinclair 2004: 127).

2.7 The justification of inductive reasoning in corpus stylistics

In 'Philosophical and literary concerns in corpus linguistics' (2011) Louw criticizes the community of corpus linguistis for focusing on being 'incorrigibly oriented towards *structure* and *description* rather than providing direct empirical meaning through the co-selection of collocates' (Louw 2011: 173). Louw claims that the vastness of empiricism supplied by reference corpora settle limitations that were encountered by philosophers. The corpus furnishes *a priori* knowledge, even of a kind that a person cannot amass in a lifetime, let alone retain and classify, and whose analysis makes it possible to explain language-related events using ***inductive reasoning***.

In Louw and Milojkovic (2016b: 263) Louw further defends inductive reasoning by quoting Mautner (2005: 654) on Whewell's ***consilience of induction*** (without induction science is impossible, Russell 1946: 647). The notion itself involves laws from several different fields being incorporated into a single theory. According to Whewell, such theories have never been found false. Louw claims that we are witnessing how Firthian

contextual linguistics, Malinowskian cultural studies, and the views of the analytic philosophers are drawn together by collocation. Louw believes that Wittgenstein could not finalise his views on necessary meaning because he could not see the QPVs and therefore did not believe that logical constants (grammar words) and their most frequent lexical collocates (metaphysics) could be representative and therefore necessary. Being frequent combinations, they represent 'the *a priori* knowledge of the impact of their grammar' (Louw and Milojkovic 2016b: 263). Philosophy largely depends on the study of ***inference***, claims Louw, and concordances facilitate inference because they 'not only lay bare the nature of events but demonstrate the extent to which outcomes are *probable'* (Louw 2017a).

It comes as less of a surprise, however, to find that collocation does satisfy Whewell's criteria for the occurrence of consilience, which are quoted in Louw (2017a) as follows:

1. When an hypothesis is capable of explaining two (or more) *known* classes of facts (or laws);

2. When an hypothesis can successfully *predict* "cases of a *kind different* from those which were contemplated in the formulation of our hypothesis."

3. When an hypothesis can successfully predict or explain the occurrence of phenomena which, on the basis of our background knowledge, we would not have expected to occur (Laudan 1981: 165, Louw's italics).

2.8 Corpus-derived subtext as prospection

In 'Teaching Literary Engagement from the Index of First lines: a Corpus-Based Subtextual Approach' (Louw 2013) the author introduces another ground-breaking empirical procedure. He claims that the subtext of the first line predicts, or ***prospects*** the development of the theme subsequently in the poem. Louw's (2013: 343) example is the poem by G. M. Hopkins, the first line of which being 'The world is charged with the grandeur of God'. Its first grammar string 'the * is *ed with the * of' yields the following subtext in the corpus of the 1995 Times newspaper:

1 ky. Forget peace and quiet, **the air is filled with** the drone of jet skis, alth

2 nd pink and shining. But the face is scored with the lines of severe intelligen
3 cted to withdraw in 1996, the sport is faced with the prospect of only 22 cars
4 tion) at Bubbua's Pit BBQ. **The air is filled with** the smell of falafel, organic
5y to be closer to 3%. Now **the air is filled with** the sound of forecasts crashing

Figure 2.1. The concordance of 'the * is *ed with the * of' in the 1995 Times corpus, taken from Louw (2013: 343).

The most frequent QPVs in the first two lexical slots, 'air' and 'filled', prospect the ending to the first stanza, as well as the concerns of the second: 'nature is never spent…'.

> And all is seared with trade; **bleared**, **smeared** with toil;
> And wears man's **smudge** and shares man's smell: the soil
> Is bare now, nor can foot feel, being shod.

The quasi-propositional variables, claims Louw (ibid: 342), are the synthetic a priori that it is now possible to reveal using corpus stylistic methods.

Prospection is also Louw's concern in 'Shared Logical Form or Shared Metaphysics: In Search of Corpus-derived Empathy in Stylistics' (Louw and Milojkovic 2015) where he goes one step further to claim that the QPVs (quasi-propositional variables) are in fact palpable descriptors of the empathy that is shared by the writer and the reader (see also Section 4.5). 'Empathy is thus seen as the sharing of the subliminal layers of meaning which a priori subsist in grammatical strings devoid of lexis, via their most frequent lexical collocates' (Louw and Milojkovic 2015: 535).

Louw explicates the procedure in philosophical terms (ibid.: 543). The sentence formed by the speaker/writer is a quasi-proposition of a possible world. It may be regarded as the given and is only a partial picture of the world it depicts. This world is amplified by the QPVs – the most frequent lexical collocates of the grammar string – and thus by the ***possible worlds*** from the reference corpus contexts. Louw for the first time brings his research in connection with the Turing test: the reply by the machine, in the form of concordances, is a sample drawn from the real world of humans and therefore passes the Turing test. The given and the reply by the computer (the corpus) share the same ***logical form***.

This assertion is based on the notion of a literary critic interrogating, say, the persona of the poem. The digital development in stylistics allows the machine to respond and produce an empirical answer to the critic's questions (Louw 2017a). 'It is collocation that sews the sentences of texts together' (Louw and Milojkovic 2015: 544).

2.9 Proximity as entailment

That 'wherever unexplained or hidden meaning is to be found, collocation as induction is likely to offer a form of instrumentation for meaning for unraveling it' is explained further in 'Entailing Proximity' (Louw 2017b). Linguistics, says Louw, lags behind other disciplines which are now considered hard science, such as physics (p. 54). He compares the discovery of corpus-derived subtext which is opaque to intuition to Max Planck's discovery of quantum that was not visible to the naked eye but was detected in measurements. In order to define the notion of ***entailment*** needed to explain the role of collocation as proximity, Louw (2017b: 55) quotes Reichenbach on the quantum: 'The philosophical problems of quantum mechanics are centered around two main issues. The first concerns the transition from causal laws to probability laws; the second concerns the interpretation of unobserved objects' (Reichenbach 1998: 1). 'If the reference corpus is viewed as a sample of the world, and the possible worlds of our creation depart from it, then we would like to measure scientifically the extent of that departure' (Milojkovic and Louw 2017: 126). In his last paper to date, 'Towards a Theory of Corpus Linguistics: Proofs Banish Proscription' (in press), Louw, for reasons stated in his previous work, claims the status of theory for collocation.

2.10 Conclusion

This section has provided an overview of Louw's published works to date. It has sketched the progress of Louw's thinking from the year 1987, when he first set in motion the discipline of corpus stylistics by comparing patterns in a poem to their reference corpus counterparts, to 2019, when Louw's Contextual Prosodic Theory may be considered a rounded scientific system. As many cross-references in this section show, Louw's thinking

has not developed linearly, but rather as continuous development of argumentation and refinement of the method.

The primary idea around which Louw's CPT has been based seems to be the notion that a balanced and representative reference corpus is a sample of its language as well as the world as represented in the language. From this it follows that the meaning and implications of any native speaker's idiosyncratic patterning may be understood in a much more nuanced way if considered from the point of view of the language norm as represented by the reference corpus, than if approached from the point of view of intuition alone. The difference between the language norm and a speaker's or writer's particular patterning Louw views as a deviation from the frequent patterns found in the corpus. As a result, Louw considers the frequent but absent collocates of a word or expression in authorial text as important as the collocates actually used by the speaker/writer.

As stated Section 1, Louw adheres to Sinclair's school of thought in corpus linguistics and considers himself a Firthian. This involves not only using untagged corpora and the bottom-up approach, and not only viewing collocation as co-occurrence ('an abstraction at the syntagmatic level', see Firth 1957: 196), but also taking into account deviation at the level of context of situation (Firth 1957: 182). It is the contextual dimension of corpus stylistic research that makes statistical collocation unsuitable for the study of meaning.

As this section has shown, research of collocation as deviation from the norm at the level of both patterning and context has taken Louw's research in two directions. One may be called 'rhetorical devices', which involve relexicalisation. The other may be termed 'insincerity' and includes not only the personal, but also the institutional and forensic dimensions. The latter involves a breach of a semantic tendency in the reference corpus that is bound to have its contextual explanation. Both directions may be explored in literary criticism, when collocation 'chunks' a writer's literary world. The deployment of absent collocates is complemented by the use of co-selection, when two expressions used in the same context yield situational contexts of a certain type in the reference corpus.

Although these developments are innovative, they still may be described as a consequence of the groundbreaking research into collocation undertaken by John Sinclair,

with the emphasis on the deviation from the norm rather than the typical. Louw's own groundbreaking discovery seems to have come with the development of subtext. In 2010 Louw postulated that grammar strings had their won 'logical semantic prosodies' that were embodied by their most frequent lexical collocates - quasi-propositional variables. A list of such variables was termed by him 'corpus-derived subtext' and claimed to shed light on the meaning of a particular authorial text, just like frequent but absent collocates would at the level of lexical collocation.

Given that corpus-derived subtext is completely opaque to intuition, Louw resorts to the support provided by the writings of analytic philosophers - Wittgenstein, Russell and Carnap. He also claims that subtext is similar to Max Planck's discovery of the quantum and that it is about time we used inductive reasoning, given that collocation has a probabilistic aspect (see also Sinclair 2004: 30). The best proof that subtext is indeed contained in discourse as part of its meaning seems to me to obtain in the mechanism of prospection, also discovered by Louw. If themes and motifs of a piece of writing are inherent in a line of text but appeal to the consciousness of the reader in the text's subsequent sections, this is empirical proof which it is not possible to deny.

In summary, by using only two methods of corpus search - co-selection and wildcarding - Contextual Prosodic Theory seems to be fully equipped to study the hidden meanings (Sinclair 2003: 117) of authorial text. The contextual dimension ensures the adequate scope of the investigation. At present, it does seem that the theory is complete and awaits its verification on the part of the academic community.

3 Semantic prosody in academic discourse, with particular emphasis on the work of Louw

3.1. The work of John McHardie Sinclair

3.1.1 Introduction

The collaboration of John Sinclair and Bill Louw lasted several decades, from 1978 (Louw, personal communication) until Sinclair's death in 2007. Sinclair was certainly Louw's greatest inspiration in corpus linguistics, and it cannot now be established with certainty how exactly their exchange of ideas, both formal (e.g. Louw's presence at the University of Birmingham as visiting scholar in 1981) and informal influenced Louw's, or Sinclair's, thinking. The dialogue was not conducted in isolation and was part of the developments within the neo-Firthian corpus linguistics community - for example, Philip (2011: 59) notes that 'semantic prosody was probably noticed and discussed informally amongst colleagues [at Birmingham University and within the COBUILD project] for some time before any of them decided to investigate it in depth'. While Louw attributes the term 'semantic prosody' to Sinclair (Louw 1993: 32), quoting it as 'personal communication 1988',[5] Sinclair (2004: 198) quotes personal communication by Louw on delexicalisation. Thus, rather than attempt to distinguish which ideas of Sinclair's influenced Louw's thinking and which originated in parallel with it, or were perhaps influenced by it, this section seeks to outline Sinclair's main views that may be either said to have shaped Louw's approach or to be able to clarify it through the congruence of these two visions.

3.1.2 The basic premises of Sinclair's approach to meaning

The basic premise of Sinclair's vision is insisting on textual evidence: 'However plausible an invented example might be, it cannot be offered as a genuine instance of language in use... In the climate of the 1970s [...] a mood of considerable humility developed, as it was realised how intricately construed each example was. Even what seemed to be innocuous changes prompted by the need to clarify a point, led to the resultant adaptation being challenged' (Sinclair 1991: 4-5). Reliance on context is insisted upon: 'the

[5] The term 'semantic prosody' was born in conversations with Louw in Harare, Zimbabwe, where Louw was affiliated with the University of Zimbabwe and where John Sinclair paid several visits as part of a link funded by the British Council (Louw, personal communication).

term *naturalness* is simply a cover term for the constraints that determine the precise relationship of any fragment of text with the surrounding text'. Form is often aligned with meaning, in fact its determiner: 'There is ultimately no distinction between form and meaning' (Sinclair 1991: 6-7). The insistence on studying natural discourse, however, entails a lexocographer's task of identifying typical, and not bizzare, instances: much depends on design criteria of a corpus (ibid.: 5; 13-26).

Primarily a lexicographer, Sinclair viewed utmost reliance on empirical data and context as key to successful reproduction of facts about language in dictionaries. Introspection was not allowed when gathering data on actual usage, but only at the stage of the evidence being evaluated by the dictionary compiler. What pertains to this thesis is not principles of compiling dictionaries, but those of text analysis. Sinclair's preferred method of gaining information on contextualised word-forms was through concordances, in the following fashion, which is also adopted by Louw: 'Quite often, it is helpful to start off with a simple KWIC [Key Word in Context] concordance and then to switch to a longer context or a sentence context for closer study' (Sinclair 1991: 43). The principle of the transfer of attention from a concordance line to a longer context seems to have eluded a number of researchers who remain convinces that neo-Firthian corpus linguistics rests on concordances alone and does not consult longer contexts in principle, i.e. that it stops where the concordance lines stop (see e.g. McIntyre 2018, discussed in Section 3.2.9).

A theory should not get in the way of verifiable interpretation, which is known as the 'bottom-up approach'. What Sinclair describes as a solution to a problem is also a practical approach that is deployed in this dissertation: 'the amount of variation in actual usage makes accurate generalisations rather difficult. The difference is often said to be between 'top-down' and 'bottom-up' approaches; starting from the 'top' it is extremely difficult to arrive at a description that fits the facts of usage, while starting from the 'bottom' it is not easy to formulate sufficiently general statements. However, the experience and intuition of the researcher are available in both approaches, and so the so-called 'bottom-up' approach properly conducted, is really a two-pronged attack on the data from the top and bottom simultaneously' (Sinclair 2003: x). A hypothesis is then modified in accordance with the more recently acquired data (ibid.: xvii). Also, corpus-driven linguistics insists on

working with very large corpora in order to obtain verifiable and replicable findings (Sinclair 2004: 193).

Another important point is Sinclair's attitude to examples that do not 'toe the line': while focusing his attention on recurrent examples, he suggests that counter-examples be studied for an explanation of their existence in the corpus: 'Language is very complex, and people use it for their own ends, without normally being conscious of the relation between their verbal behaviour and the way that behaviour is characterised. They are creative, or expedient, or casual, or confused; or they have unusual matters to put into unusual words, so they have to combine them in unusual ways' (Sinclair 1991: 101). He realises that oddities are there for a reason. Among other things, the reason can be a historical change under way (ibid.: 61), but it is important that 'no instances are overlooked in any sample, no matter how awkward or bizarre they might be (ibid.: 94).

Sinclair is also aware of the difficulty faced by the corpus analyst, voiced in Louw and Milojkovic (2016: xvi) - that in theory corpus linguistic analysis of corpus data never ends: 'no matter how thorough and comprehensive a study is, there is always a lot left over for an even more detailed study' (Sinclair 1991: 65). Also, Sinclair's findings related to the interaction between lexis and syntax point to Louw's (2010a, 2010b) discovery of corpus-derived subtext (the most frequent lexical collocates of a grammar string) and how it can influence text interpretation: 'It seems that there is a strong tendency for sense and syntax to be associated [...B]y using the same evidence in reverse, the traditional domain of syntax will be invaded by lexical hordes' (Sinclair 1991: 65). He proposes to broaden the area of syntax to include lexical structure as well (ibid.: 104). 'A grammar is a grammar of meanings and not of words. Grammars which make statements about undifferentiated words and phrases leave the user with the problem of deciding which of the meanings of the words and phrases are appropriate to the grammatical statement' (Sinclair 2004: 18).

The interconnectedness of the segment of text under consideration, the text and the world, and the text and authorial intention, all of which Sinclair refers to, is very much in the spirit of Louw's subsequent discovering proof for his Contextual Prosodic Theory in the work of Wittgenstein (see Section 2): 'Such is the intricate nature of the ties between one segment of text and the surrounding text, and the relation between the text and the world

and the intended outcomes of communication, that the act of plucking a few words from any text is not likely to provide a freestanding instance of its constituent words, each acting typically' (Sinclair 1991: 99). Examples that are typical will thus have to be found by statistical means. The word in the centre of the concordance is called the ***node***, and the word-forms that appear in it are called its ***collocates***.[6] Statistically, the node's attraction lessens considerably beyond four words from the node.[7] The researcher determines the suitable cut-off point (the author offers an example of one-tenth of the frequency list, and the concordance is sorted in the order of frequency, i.e. typicality (ibid.: 105-106).

Separating lexis from syntax would be erroneous, as most everyday words only have meanings when they are components of larger structures. Perceiving language as abstract syntactic patterns with lexical slots (the open-choice principle) is possible in some special cases. The majority of meanings is formed by frequent words as part of multi-word patterns, despite superficial variation (the idiom principle) (Sinclair 1991: 109-110). Idiomaticity becomes at least as important as syntax when it comes to the construction of meaning (Carter 2004). The idiom principle underlies structural tendencies. These same tendencies are in operation in grammar strings which attract some lexical variables more often than others, and this mechanism underlies the existence of corpus-derived subtext. In connection with this, Sinclair affirms that there is very little reason to posit a semantics independent of corpus studies, if the purpose is text description (Sinclair 1991: 121).

It is important to note that the boundaries of meaning are not fixed but fuzzy: 'One way of describing collocation is to say that the choice of one word conditions the choice of the next, and of the next again [...] Successive meanings can be discerned in the text, and you can associate a meaning or a component of meaning or a shade of meaning with this or that word or phrase that is present in the text. But it is often impossible in the present state of our knowledge to say precisely where the realisation of that meaning starts and stops, or

[6] This is not how collocation is defined by Baker, Hardie and McEnery (2006: 35-36): '*Collocation*, according to Baker, Hardie and McEnery (2006), refers to the phenomenon that particular words have a higher likelihood of occurring in association with other certain words in particular contexts. Therefore, a collocate is a word which occurs with frequency 'within the neighbourhood of another word' (Alsuweed 2015: 24).

[7] In Russian, due to its inflectional morphology, the attractive power of the node has been found to be 6, and not 4 words from it (I. Azarova, personal communication).

exactly which pattern of morphemes is responsible for it [...] In the way in which we currently see language text it is not obvious how each small unit of form prospects the next one (Sinclair 2004: 19-20).

3.1.3 Collocation, colligation, semantic preference, semantic prosody

Sinclair (2003: 145-6) distinguishes between collocation (the co-occurrence of particular words with the node), colligation (the co-occurrence of a grammatical class with the node), semantic preference (the co-occurrence of lexis belonging to a particular semantic set or sets with the node) and semantic prosody (the specific tendency of meaning that is formed in the reference corpus by the node and its collocates). According to Sinclair, this is a cline of abstraction of categories which can describe a lexical item, starting with collocation as the most concrete and finishing with semantic prosody as the most abstract. For example, the expression 'true feelings' is normally preceded by a possessive adjective or the definite article (colligation), which is in turn preceded by a verb denoting expressing or stifling one's true feelings, such as 'reveal' or 'deny' (semantic preference), and the element of reluctance or difficulty that is habitually part of the meaning of a collocation featuring 'true feelings' is this expression's semantic prosody. This clearly shows the overriding importance of semantic prosody: the semantic field that is part of semantic preference is a structural tendency but it alone does not determine meaning (rather, it may be useful in language description and teaching). On the other hand, the meaning of the lexical item cannot be described without resorting to semantic prosody, because it is in fact the point of using the lexical item in the first place (ibid.: 148), which is why Sinclair posits that semantic prosody is in fact the only obligatory element of a lexical item, apart from its invariant core, responsible for the communicative effect (Sinclair 2004: 122, 141, 174). It is 'a kind of meaning that is subliminal, in that we only become aware of it when we see a large number of typical instances put together' (ibid.: 18). Thus Sinclair affirms that apparently problematic link between what is factual and what is subliminal: '[F]acts of English that lie below the surface, usually below consciousness, have their effect on meaning but are difficult to pin down without external evidence and appropriate tools' (ibid.: 127). Also, Sinclair acknowledges that meaning is always unique: '[S]ome aspects of

textual meaning arise from the particular combinations of choices at one place in the text, and there is no place in the lexicon-grammar model where such meanings can be assigned. Since there is no limit to the possible combinations of words in texts, no amount of documentation or ingenuity will enable our present lexicons to rise to the job. They are doomed' (ibid.: 134).

3.1.4 Sinclair on prospection

Prospection, to Sinclair, is linked to the probabilistic aspect the idiom principle and is, through phraseology, connected to the notion of repeated events: if a word is not completely independent in its patterning because it is very rare or has a protected status (e.g. technical words), it will be part of collocational patterning that will follow language users' phraseological expectations, which are simultaneously 'expectations of events as collocations arise' (Sinclair 2004: 30). Sinclair's focus here is on the regularity of patterning, but the mention of repeated events that are, because they are repeated, reflected in predictable usage of certain words is also the basis of Louw's corpus stylistics approach: the patterns in authorial text need preferably to be studied in those contexts in the reference corpus that contain similar events (see e.g. Section 2.6.2.2). Sinclair claims that '[l]anguage models experience and is adaptable enough for the job, but the price to be paid is representation through a largely prospecting organisation' (2004: 53).

3.1.5 The word vs. a unit of meaning

Sinclair's focus is not on the word (as defined somewhat indecisively in traditional terms in Sinclair 1991: 176) as the starting point of description of meaning, but on extended units of meaning: 'So strong are the co-occurrence tendencies of words, word classes, meanings and attitudes that we must widen our horizons and expect the units of meaning to be much more extensive and varied than is seen in a single word' (Sinclair 2004: 39). Semantic prosody is seen to play a leading role in the integration of an item with its surroundings. 'The initial choice of semantic prosody is the functional choice which links meaning to purpose; all subsequent choices within the lexical item relate back to the prosody' (ibid.: 34). The structural categories of a lexical item (collocation, colligation,

semantic preference, semantic prosody) may assume a central and not peripheral role in language description, argues Sinclair (2004: 39). This is in agreement with Louw's view of collocation as instrumentation for meaning. Interestingly, the very last paragraph ever to be written by Sinclair was on semantic prosody (Sinclair 2010:45).

3.1.6 Conclusion

This section has dwelled on the theoretical premises which John Sinclair, the originator of the Birmingham school of corpus linguistics and the COBUILD project, shared with Bill Louw. These premises are characteristic of the corpus-driven bottom-up approach generally speaking - unbiased reliance on raw untagged corpus-data, on the one hand, and viewing phraseology as probabilistic when it comes to recipients' expectations, as well as reflecting and bearing traces of repeated events in the surrounding world and the experience shared by the participants of the communicative situation. In particular, when considering the structure of a lexical unit, Sinclair attached major importance to semantic prosody, which, as a derivative of collocation, and together with delexicalisation/relexicalisation (as was shown in Section 2.2.1) was a starting point in the development of Louw's Contextual Prosodic Theory prior to the introduction of logical semantic prosody - subtext. Sinclair's considering SP an obligatory element of a lexical item, together with its core, highlights the link between the theoretical approaches of these two scholars, especially when it comes to Louw's views on collocation as instrumentation for meaning. The next section will dwell on the modifications and further adjustments that other scholars have brought or sought to bring to the 'concept', as some of them called it, of semantic prosody, together with some criticisms of Louw's approach as they appear to have understood it from his writing.

3.2. The perception of semantic prosody in corpus linguistics

Particularly since the publication of Louw's 'Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies' (1993), the phenomenon of SP has attracted a great deal of interest. This is understandable, given the properties of semantic prosody that had been highlighted by Sinclair and Louw, such as its being an obligatory element of a lexical unit, its subliminal nature and especially the tricks it plays on meaning

when the language user is in breach of the language norm. Notable progress in the field was made by Stubbs (1995, 2001), Tognini-Bonelli (2001), Hunston (2002, 2007, 2011), Partington (2004, 2014), Forest (2007), Bednarek (2008), Morley and Partington (2009). Semantic prosodies have been examined cross-linguistically and from the point of view of translation (Berber-Sardinha 2000, Xiao and McEnery 2006, Dam-Jensen and Zethsen 2008, Stewart 2009, Kübler and Volanschi 2012, Ebeling 2013, Wei and Li 2014), language teaching (McGee 2012, Fuqua 2014), cognitive studies (Oster 2010), as well as experimental psychology (Hauser and Schwartz 2016, 2018).

However, the academic discussion surrounding the notion of semantic prosody has shown several specificities, particularly with reference to the work of Louw. Despite the fact that studies of 'set in' and 'happen' get a mention in Sinclair as early as in 1991 (pp. 73-75, 112), and despite Louw's attribution of the coinage of the term to Sinclair (Louw 1993: 32), in much of the literature semantic prosodies as a phenomenon are associated primarily with Louw (e.g. see McEnery and Hardie 2012: 135, Williams 2010: 403). Interestingly, several attempts are made to modify the term itself: Stubbs (2001: 65) suggests 'discourse prosody', O'Halloran (2007b: 16) introduces 'register prosody', Hoey (2005: 24) opts for 'semantic association' vs. 'pragmatic association', Hunston (2007: 266) proposes 'semantic preference'/'attitudinal preference' vs. 'discourse function', and Partington (2014) opts for 'evaluative prosody'. Accounts of what is perceived by researchers as Louw's contribution range from sympathetic (Philip 2011) to highly unfavourable and exhibiting unnecessary bias (McEnery and Hardie 2012).

Another feature that characterises this discussion is that, despite the steady flow of publications by Louw after 2000, there are very few mentions of Louw's work published between 2000 and 2010 in the area of corpus studies, and a complete dearth of commentary on corpus-derived subtext (Louw 2010a, 2010b). The attraction of showering criticism on Louw (1993, 2000) to the point of completely ignoring his later work is such that the most recent excursion into semantic prosody attempted by McIntyre (2018) still dissects these early publications, after twenty-five and eighteen years respectively! Consequently, no discussion of a reaction in the field of corpus linguistics to Louw's CPT as a whole can be included in this dissertation, except for the mention of Wang and Humblé (2017) in the next

section. Hence this section is devoted primarily to studies of corpus linguists where Louw's work either comes under attack, or where the phenomenon of semantic prosody receives detailed treatment that helps further to clarify its role in discourse and meaning construction and construal.

One of the difficulties of describing the academic discussion around the issues related to semantic prosody is that each author attempts to review the situation in the field, offers their own interpretation of it, and quotes everyone else involved in SP. As a result, there exists a circle of authors quoting one another and debating with one another, in addition to quoting Louw and Sinclair. Therefore, in the attempt to avoid confusion, the order of appearance in this review will be for the most part chronological, unless circumstances dictate otherwise.

3.2.1 Refinement and re-naming of semantic prosody

At first sight, considerable modifications to the views of Louw and Sinclair are offered by Hoey's theory of priming. As explained in Louw and Milojkovic (2015), its major difference from these scholars' views is Hoey's taking Louw's empirically verifiable semantic prosody into the domain of mental concordances, 'richly glossed for social, physical, discoursal, generic, and interpersonal context' (Hoey 2005: 11). Replacing the reference corpus (with its concordances) with its mental counterpart, presumably specific to each individual, Hoey enters the realm of the psychological. He prefers the notion of 'semantic association' to those of semantic prosody or semantic preference, while reserving the term 'pragmatic association' for the discourse function of SP, 'when a word or word sequence is associated with a set of features that all serve the same or similar pragmatic functions (e.g. indicating vagueness, uncertainty)' (ibid.: 26). Stewart (2010: 153) notes that Hoey (2005: 157) himself admits that this description corresponds to Sinclair's view of semantic prosody. 'Priming prosody' is not a replica of 'semantic prosody', but echoes Sinclair's description of the lexical item with the added textual dimension (Hoey 2005: 166). The author also introduces 'priming conflict' which occurs 'when a choice of one priming is overwhelmed by another, more dominant priming. The result here is either ambiguity or humour' (ibid.: 170).

For similar reasons, i.e. in order to highlight the pragmatic function of semantic prosody, as well as its cohesive function in discourse, semantic prosody had already been given a new name. In his influential study on the negative prosody of the lemma 'cause', Stubbs (1995) compares it with 'create', 'affect' and 'happen', and also with 'reason', 'result', 'effect' and 'consequence'. He advocates cautious use of quantitative methods, combined with the analyst's introspection. In a larger-scale study, taking stock of semantic prosodies of a great number of lexical items, such as 'accost', 'heritage' and 'loiter', Stubbs (2001) proposes that this phenomenon be referred to as 'discourse prosody', 'to emphasize their [prodosies'] function in creating discourse coherence (Stubbs 2001: 65–66).

O'Halloran (2007b) accepts the role of semantic prosody in stylistic analysis, but also sees fit to introduce the term 'register prosody' with the following justification (here discussing the semantic behaviour of 'simmer'): 'Although there are only a few examples from fiction, nevertheless the positive associations for the past tense (with or without progressive aspect) clearly contrast with the negative associations of the perfect progressive use, "auxiliary+been+simmering", in hard news and neutral associations in recipes' (O'Halloran 2007b: 16). When it comes to the contexts of fiction in the Bank of English, the author adds that 'was simmering' might turn out to be predominantly positively used in romantic fiction generally, i.e. that it had a genre prosody. Thus, the author introduces a further categorisation of semantic prosody, adding the categories of 'register prosody' and 'genre prosody'.

3.2.2 Partington (2004)

One of the influential studies in the field of semantic prosody is Partington's (2004), who is concerned with the evaluative aspect of semantic prosody. The evaluative meaning, mainly taken as good or bad, is not inherent in a word (he quotes 'timely' and 'flabby' as examples of built-in meanings), but 'potentially goes beyond a single orthographic word and is much less visible to the naked eye' (Partington 2004: 132). In his detailed study of 'happen' the author also notes the text-cohesive and semi-modal function of the verb, expressing a lack of certainty or absence of factuality (ibid.: 137-139). The author also points to Stubbs's accounts of the inter-relatedness between the negative prosodies of

'accost', 'loiter' and 'lurk' and their syntactic structures (Stubbs 2001: 198-206). Following Sinclair's (1996: 91) discussion of 'brook' and Stubbs 's (2001: 66, 89-95) discussion of 'cause' and 'undergo', the author also concedes that semantic preference is a 'narrower' phenomenon contributing to building semantic prosody (Partington 2004: 151). In this, he is supported by Bednarek (2008).

3.2.3 Whitsitt (2005)

Whitsitt (2005) appears not to share Louw's view that an awareness of semantic prosody is an improvement on the traditional concept of connotation. He seems to disregard the important difference that in the case of semantic prosody the centre of the investigation is the node, and in the case of connotation the centre is a mental concept. Moreover, semantic prosodies are a mechanism that can provide an empirical account of connotations, which is, after all, Louw's point.

Then, Whitsitt goes so far as to suggest that the 'concept' of semantic prosody should be dismissed altogether because Louw (1993: 37) in passing pronounces them 'undoubtedly the product of a long period of refinement through historical change' without offering proof and while focusing on a reference corpus which is synchronic and not diachronic (Whitsitt 2005: 287-288).[8] This objection is refuted by Morley and Partington (2009: 152, 156). They point to research done in the area of historical linguistics, for which pejoration and amelioration, narrowing and broadening of meaning, and metaphorisation are central notions.

The analogy with the phonological sense of prosody Whitsitt also finds disputable - sounds do not get permanently affected by proximity with other sounds the way nodes are by proximity with their collocates (Whitsitt 2005: 291). Louw gets to blame for this as well, although the term 'prosody' has never been subject to alteration in the literature, unlike its modifying adjective. More importantly, this argument is also faulty. It is true that /k/ sounds different depending whether it is followed by a front or a back vowel, but these phonological changes are consistent in much the same way in which words change their

[8] Louw and Milojkovic (2016: 178) provide an illustration of how the semantic prosody of 'being what it is' has changed over time. In any case, it is not diachronic change but the empirical verifiability of semantic auras that is central to SP.

implications depending on their co-text. To sum up, the author states that 'Louw's arguments for semantic prosodies are based on metaphors which are not convincing, analogies which do not hold up, and no empirical evidence' (Whitsitt 2005: 296). It can be concluded from this that Whitsitt considers his own statements both convincing and well-argumented.

3.2.4 Hunston (2007)

Hunston's analysis of 'to the point of' (Hunston 2007) is also an instance of corpus skepticism, but differently motivated: it has to do with the irony/insincerity dichotomy proposed by Louw (1993). The expression 'to the point of' is used negatively in an overwhelming majority of cases. However, the author quotes an example where it is atypically followed by a positive collocate without any indication in the rest of the text that there must be either irony or insincerity involved. Hunston emphasizes that 'explanatory' power of the corpus (which she accepts) must not be confused with 'predictive' (which she intends to disprove by this example). It cannot be proved, she argues, that the author was lying or joking, or that all readers will assume she was (Hunston 2011: 66 also adheres to this view). Milojkovic (2013a) argues that in the case of insincerity other context clues may very well be absent. Insincerity must not be understood only as an attempt at lying. In the majority of cases we will modify our very private views somewhat when communicating them to others. The journalist in Hunston's example, who used 'to the point of' atypically, expressed the view she intended to express in her article. However, the reference corpus encourages the hypothesis that she personally did not whole-heartedly support the view she was expressing. Writing an article in order to make a living is not the same as writing a poem, for example. Louw's point is that a stylistician armed with a corpus will give an empirical account of the journalist's reservation, whereas the reader's intuition will not.

Also, Hunston concludes that ascribing semantic prosody to a word is oversimplistic (p. 256) and is inclined to follow closely Sinclair's 'more subtle' view of SP as 'the consistent discourse function of the unit formed by a series of co-occurrences' (p. 257).

3.2.5 Morley and Partington (2009)

Hunston's criticisms are, to my mind, aptly dealt with in Morley and Partington (2009). Firstly, the authors emphasise the cohesive function of semantic prosody, as it in effect creates a consistency of evaluation across words. They believe SP to be 'a fundamental aspect of the idiom or phraseological principle of language production and interpretation, in particular providing convincing evidence of how elements of meaning 'hunt in packs' (this is an allusion to Louw (2000) and is here to underline the consistency with which an evaluative meaning is maintained in a stretch of text through the use of successive semantic prosodies). The authors are convinced that semantic prosodies can mostly be categorised as either good or bad, but that they also express specific meanings (Morley and Partington 2009: 141-142).

Turning to Hunston's (2007: 258) preference for viewing semantic prosody as belonging to a unit of meaning rather than to a word (which is another area where she sees herself as disagreeing with Partington 2004), the authors claim that this is a false dichotomy. On the one hand, words (or lexical items, as this term includes multi-word units) do carry a set of instructions on how to use them (this explanation is in accordance with Hoey's (2005) theory of priming; Louw would probably say that a word is a corpus lingustic starting point that can be investigated in a reference corpus). On the other hand, from the discourse perspective, semantic prosodies are 'definable as the evaluative intent of the speaker [...] in any individual text fragment' (p. 147); from the corpus-attested or statistical prospective, the item will have an invariant core in different patterns of co-occurrence, allowing us to study the default prosody and counterexamples (p. 148; see also Sinclair 2004: 30-35, 114; Sinclair 2010: 45).

3.2.6 Stewart (2010)

Stewart (2010) attempts a comprehensive overview of the literature on semantic prosody, highlighting and clarifying issues that, to him, seem to surround what he calls the 'concept' of semantic prosody (Stewart 2010:1). This is the first difficulty the reader encounters - is semantic prosody the property of a lexical item (or a stretch of text) or a 'concept'? The explanation might be that Stewart's goal is not investigating semantic

prosody as an empirical phenomenon, existing in the language separately from what opinions people may have formed on the subject, but precisely trying to systematise, respond to and reconcile the existing opinions and ideas surrounding what ultimately does come across as a notion from different researchers' perspectives.

What I consider the drawback of the whole volume (whether the topic is the synchronic or the diachronic aspect) is Stewart's main focus on co-text, i.e. habitual co-occurrences, not the context of situation in the Firthian tradition, despite his detailed account of Louw (2000), where it is given primacy (Stewart 2010: 13-14). To my mind, this would have freed the author of much of his general attitude of skepticism. For example, Stewart writes:

> [...] consider the following two occurrences of *CAUSE* as verb in the BNC:
>
> - The door closed and Elaine pulled the magazine in, causing the letter box to snap shut smartly.
> - The inhibitors might therefore be causing amnesia not because they prevent protein synthesis but because of their effect on increasing animo acid levels.
>
> Here once again it seems problematic to postulate that some sort of attitude is being expressed, because it is not clear of whom or of what the respective speakers/writers might disapprove. Of letter boxes? Of Elaine? Of inhibitors? (Stewart 2010: 23)

Even a superficial stylistic analysis would be sufficient to explain that 'cause' in the first passage does not refer to Elaine, but to the verb 'snap', which in this context contributes to the reader's overall impression that Elaine's movements are definite, brisk, energetic (the latter judgment is confirmed by the adverb 'smartly'). The postulated negativity of 'caused' here may refer to the sound: 'If something snaps shut or is snapped shut, it closes quickly with a sudden sharp sound'.[9] As for amnesia, not many people would refer to it approvingly (not to mention the undisputed fact that 'cause' does not carry disapproving overtones in scientific discourse, as stated in Hunston (2007: 263)).

[9] https://dictionary.cambridge.org/dictionary/english/snap

Reluctance on the part of Stewart to take into account Firthian situational context leads to claims such as that words like 'alleviate' and 'relieve' are also commonly followed by negative collocates, in the same manner as 'cause' is, and yet their positive implications seem to be untainted by them![10] Louw (2000) deals substantially enough at least with this observation by introducing the term *fractured context of situation,* whether positive or negative, and this innovation does not escape Stewart (Stewart 2010: 13). Why he insists than an SP is wholly dependent on the node's first collocate to the right, remains unclear.

Taking things too far can best be observed in Stewart's discussion of Louw's view on prosodic clashes. Taking 'a clash' to mean 'an unusual combination', or 'any word combination that does not exist in the corpus' (instead of acknowledging it to be an evaluative usage goes against the corpus norm), Stewart in all seriousness discusses Louw's collocation 'semantic aura' as an unusual combination which is neither ironic nor insincere and in need of a third explanation (Stewart 2010: 25). Also, the author seems to believe that Sinclair's and Louw's views on SP differ when it comes to determining the length of the unit that is characterised by a prosody. While Sinclair attributes it to the entire lexical item, such as 'true feelings', Louw is reported to believe that semantic prosody, though dependent on the co-text, is the property of a word (ibid.: 57, 59, 161; McEnery and Hardie 2012: 136). He obviously ignores Louw's examples such as 'fine friends' or 'days are', discussed in the previous section.

Despite these objections (which may be summed up, in my view, as Stewart's taking too literal an approach to some of the claims encountered in the literature), the book is a valiant attempt at clarifying a number of assertions in the literature and thus helpful in arriving at an unassailable definition of semantic prosody. Firstly, he quite rightly doubts that native speakers are all that incapable of discerning semantic prosodies through introspection, and states correctly that no detailed studies on this score have been carried out (Stewart 2010: 81-82).[11] Then, Stewart states that the word 'collocate' in Louw's papers

[10] This apparent paradox is solved by Partington (2004: 155): if the relationship between collocates is that of opposition or detraction, he claimes, the node is not affected by the negativity of the sum of its collocates.

[11] My academic writing class, in their fourth year of English studies at the University of Belgrade, guessed both the SP of '*brook* 'and *'true feelings'*. When asking about the latter, however, I judged it fair to adjust the intonation (indicative of a position in a sentence in progress) for better results.

refers to a co-occurring word regardless of the number of occurrences, and not to a word frequently co-occurring with the node, which is how it may be understood by others, especially since he seems to understand the word 'collocation' as habitual co-occurrence (Stewart 2010: 86-87). A semantic prosody may be established on the basis of very many words sharing the same semantic feature or features even if many among them appear in the concordance only once. More importantly perhaps, Stewart draws attention to a certain ambivalence on the part of various researchers when they interpret concordances they show in print.

If the researchers quoted by Stewart might on occasion be accused of haste and arbitrary judgment, as he himself fears (Stewart 2010: 103), his own view of the concordance as an instrument of research is nothing but puzzling. The concordance, he claims, 'becomes a toil, above all because we are simply not accustomed to reading text in this way. We are used to the traditional method of repeatedly reading from left to right and then down, and we are not used to seeing sentences lopped at the beginning and end of the line, nor to seeing the same word occurring repeatedly down the middle of the page' (Stewart 2010: 105). All in all, far from rejecting the concordance as a tool of corpus research, it could be argued that Stewarts's account has been helpful in listing all possible obstacles that a novice in corpus research could encounter (some of them supported by Section 12 of this thesis).

3.2.7 McEnery and Hardie (2012)

Hunston's view of the distinction between the explanatory and the predictive potential is supported by McEnery and Hardy (2012), who proceed to express still more skepticism. Milojkovic (2013a) focuses on their arguments one by one. Talking about Louw's example of irony, 'President Clinton fanned the flames of optimism in Northern Ireland' (see Section 2.2.3.1), the authors state their first objection: 'It is trivially easy to demonstrate that the reproduction of a fixed linguistic pattern whose meaning runs contrary to the speaker's own expressed opinion is not evidence that the opinion is expressed insincerely. For instance, an English-speaking atheist is not unlikely to exclaim *Oh my*

God! when astonished or *Bless you!* when another person sneezes. This does not mean that the espousal of atheism is insincere' (McEnery and Hardy 2012: 140).

It is not relevant to corpus stylistics analysis whether the speaker believes in God. The expressions given here as examples are used delexically, as is shown in the last 20 lines from the concordance of 'Oh my God' (in line 35 ‘God’ is purposefully relexicalised):

```
20  his composure at once, crying: "Oh my God, Nicole is killed? Oh my God, she’s
21 their head in the sofa, moaning: "Oh my God!" No interviewee ever treated Carl
22  my mother was outside shouting, ‘Oh my God, oh my God’." Only once in Millig
23 was outside shouting, ‘Oh my God, oh my God’." Only once in Milligan’s lifeti
24 burn said: "Mrs Homer screamed, ‘Oh my God, please leave him alone’, at which
25 t know who that is." I explain. "Oh my God," says Jonathan, looking shocked. "
26 ling her the date and the time. "Oh my God," she replied. She had thought it w
27 ing: "Oh my God, Nicole is killed? Oh my God, she’s dead!" Marcia Clark, for t
28 ‘m really scared," and I replied, "Oh my God, so am I!" Eventually, the chaps
29 es, I’d whipped me kit off and, ‘Oh My God someone’s having fun in opera’. Hor
30 er their dishevelled apartment ("Oh my God, the Kandinsky!", Ouisa shrieks), G
31 ionOfPaper> <Story> <Group> Oh my God. The roistering royals are at it ag
32 inviolable novel stuff, all the "Oh my God, they’re not snogging in crinolines
33 ke Renault, Ford has not thought: "Oh my God. We’re going to charge 20 grand f
34 ses the all too common blasphemy, "Oh my God!" What does it say about our soci
35 working on it. Mary blurts out "Oh my God!" when confronted by the Angel Gabr
36 sore ..." followed by "1,2,3,4 ... Oh my God, where’s my logic gone?" Coma is
37 ("Uff! Oof!") finished him off. "Oh my God," yelled the cameraman, and bolted.
38 t’s terrific, it’s so big, it’s Oh my God ... yes!" Has nobody told Evans that
39 , it’s me, "I said desperately. "Oh my God, you look like my mother," she said
40 at home listening to music. She said, ’Oh my God, you’re going to hurt your vo
```

Figure 3.1 Concordance of 'Oh my God' from the 1995 *Times* corpus

The basis of Louw’s approach is studying deviations from the norm, as represented by a large and balanced reference corpus. It is irrelevant whether the speaker who has used a delexical expression believes in the expression’s fully relexicalised meaning. Both ‘Oh my God’ and ‘Bless you’ have so long been used delexically that the former, when relexicalised, is the basis of humour in a BBC Christmas radio show described in the wider context of line 35.

The authors then explain that 'fan the flames of' was found by them to be negative in the reference corpus, bar two examples. Milojkovic (2013a) analyses these two examples in detail and proves that by taking into account Firthian context of situation and Malinowskian context of culture, these two occurrences are also instantiations of a negative SP. It would appear that McEnery and Hardy (2012) regard all interpretation as personal and subjective, and therefore claim that it is unfalsifiable, as well as that semantic prosody cannot be used

predictively. On the contrary, a corpus stylistic analysis is *replicable* because it is *verifiable* through large reference corpora and embedded in the Firthian-Malinovskian context of situation through *similar events.* Context clues are an objective factor; their interpretation is not individual or intuitive but corpus-based; a reference corpus is the accumulated experience of the language norm; similar events, found so within the framework of the context of situation, ensure a line of best fit in verifiable and replicable interpretation. When corpora have grown much bigger, the findings of this thesis may be *falsified.*

3.2.8 McIntyre (2018)

Perhaps the most interesting contribution to the discussion on SP is that by McIntyre (2018). It is interesting for two reasons. Firstly, the author, after providing a general overhaul of the state of affairs concerning views on SP in the literature, offers new objections to Louw (1993; 2000), which, to my knowledge, have not been raised elsewhere. Secondly, he actually tackles the question of irony, first raised by Louw (1993), and provides his own analysis of a comic sketch by comedian Alan Bennett on the topic of World War II, featuring this discoursal tactic:

> Then America and Russia asked if they could join in, and the whole thing turned into a free-for-all. And so, unavoidably, came peace, putting an end to organised war as we knew it (Bennett et al. 1987: 78).

For the latter enterprise, he claims that a 'more nuanced approach to semantic prosody' is necessary, as well as utilising semantic preference and Gricean maxims. In particular, McIntyre claims that irony in the statement by the comedian 'so, unavoidably, came peace' cannot be established on the basis of SP alone, but such an analysis must be assisted by Gricean maxims of quality and manner. According to him, the maxim of manner is flouted by the combination of 'peace' and 'unavoidably'. Besides, the statement that peace 'came' flouts Gricean maxim of quality because peace is normally brought about by human efforts, so the statement must be false.

This whole analysis is objectionable at several levels. First of all, it is unclear why the author should resort to Gricean maxims (Grice 1975) in the first place. These maxims originated as an attempt to prove that there are no divergences between formal devices in philosophical logic and their counterparts in natural language. Grice introduces a 'rough general principle', which he calls the Cooperative Principle, presupposing that the participants have entered conversation in good faith and want it to succeed. He assumes that there are four categories (or maxims) of this principle which ensure successful communication: Quantity, Quality, Relation and Manner.

The collocation of 'unavoidably' and 'peace' is definitely odd in the sense of its being unusual, as McIntyre correctly says, but that does not make it unclear. The Gricean maxim of manner warns against cases of obscurity, ambiguity, prolixity (verbosity) and lack of order. The researcher is not helped by the collocation of 'unavoidably' and 'peace' being categorised as obscure, ambiguous, or incoherent, because it is none of the above - it is ironic, no more, no less. It is not the corpus that diagnoses irony, it is the investigator, and he or she does so after juxtaposing corpus data and the context of situation. Interestingly, irony is categorised by Grice (1975) as a breach of the maxim of quality - therefore, as insincerity with a conversational implicature resulting in an ironic effect. This link between irony and insincerity inadvertently supports the main premise of Louw (1993).

The same applies to McIntyre's conclusion that the collocation 'came peace' is in breach of the maxim of quality because it is false as peace does not come unassisted by human endeavour. This is not a valid interpretation. Peace treaties are indeed signed by human agents, but never has it been a privilege of the ordinary people, who, after all, create the language norm - they eagerly await the coming of peace but have no influence over it. On the other hand, the BNC concordance of 'came' followed by a noun (the noun is a non-human agent, as in the text under investigation; 2 is the cut-off point) shows that what comes is indicative of a significant transition which may be fortunate or not:

came time 7
came news 5
came death 3

came reports 3

came anger 2

came number 2

came sounds 2

came thunder 2

Co-occurring with 'unavoidably', 'peace' falls into the category of negative transitions, which intensifies the subtle irony in Allan's wording. It seems that Gricean maxims, being an intuitive attempt to provide a philosophical foundation for the phenomenon of conversational implicature, do not refine a corpus stylistic analysis, but may be refined by it.

3.2.9 Toolan's views on semantic prosody and his treatment of prospection

3.2.9.1 Toolan (2007)

Toolan (2007) discusses Sinclair's *Trust the Text* (2004) 'in which several radically new ideas about the role of prospection and encapsulation are outlined' (Toolan 2007: 269). To Toolan, the notion of semantic prosody is inherent in the very definition of collocation: 'collocation is the marked tendency of words to co-occur in patterns or structures where, by virtue of the frequency of the association, part of the meaning of the pattern is distributed like a prosody across the entirety of the pattern and cannot be assigned exclusively to one or another of the component words (like opaque idioms, the pattern also resists semantic decomposition)' (Toolan 2007: 272). Expressing particular interest in the prospective features of discourse and how they can be revealed by corpus means, Toolan nevertheless expresses several reservations. Studying a novel or a short story, he believes, must take into account the text's completeness (the rounded plot, the message conveyed, the characters' development) , in contrast to non-literary genres where there is no such final quality. This reservation is countered by Simpson's (2014:48) description of corpus stylistics as a discipline studying authorial usage as deviation from the corpus reference norm (see section 3.3.2).

Despite the overall praising tone, Toolan considers Sinclair's *Trust the Text* to be work in progress and in need of future verification. He also mentions, towards the end, that

we should not 'overstate the strength of particular phraseological collocations, and even of the phraseological tendency [...] such as the alleged negative prosody of "weather" and "set in" (Toolan 2007: 285). In the BNC and COCA there is only one example of weather+set* in, which might be thought of as non-negative per se:

We suspect that your golden elder is basically quite healthy but, as soon as the warmer, drier weather sets in, it is showing signs of drought stress (The Gardener. London: Maxwell Consumer Magazines, 1991).

Although some readers might think of 'warmer, drier weather' as of a pleasant change, the plant in the focus of the sentence seems by no means to enjoy it. This is an adequate reply to Hunston's (2007) and other authors' concern for the point of view in establishing the positivity or negativity of semantic prosody, and to Toolan's 'what if' question. Furthermore, Toolan's objection (he is a native speaker of English) can serve as proof of the inability of intuition to access semantic prosodies, which Louw calls 'intuitive opacity' (Louw in press).

3.2.9.2 Toolan's work on prospection

Prospection and readers' expectations are in the focus of Toolan's (2009) semi-automatic corpus stylistic approach. Toolan's first step in the analyses is based on trial and error - he notes, for example, that certain patterns in the short stories he studies can form a kind of skeleton that outlines the plot of the given short story. He finds that these searchable patterns can be key words (especially if these are the names of protagonists collocating with action/dynamic narrative-tense verbs), fully lexical key words and clusters, first sentences of narrative paragraphs (not just any paragraphs), sentences containing the protagonists' represented thought or direct speech, questions and requests, negation-carrying sentences, and sentences containing narrative verbs of modality and mental processing.

His initial exploratory steps in finding 'textual presences... [which] guide the reader to a sense of plotted and expected continuation' (Toolan 2009: 136) are in their majority based on his linguistic experience outside corpus techniques, and are then checked

empirically on a particular sample. In other words, the choice of his parameters is based on introspection, the method of confirming their prospecting qualities is empirical, and the subsequent search for them in the short story is automatic, or semi-automatic.

Toolan's understanding of prospection is founded on Sinclair's (2004: 82-101, 115-127) premise that a textual segment may 'raise implicit questions that subsequent discourse will address, wholly or in part, and that each succeeding segment of text may open or prospect further kinds of question or incompleteness for yet later text to resolve' (Toolan 2009: 16). 'The main textual resources in the modern short story which […] are particularly instrumental in guiding the reader's expectations' he calls 'narrative expectation' from the point of view of the reader, and this same phenomenon he terms 'prospection' if looking at it 'as the elements of text design that cause expectation' (ibid.: 114). Thus, the author does not set out to discover what predicts the outcome, only what generates expectations - that can later be modified, flouted or fulfilled.

On the other hand, Louw's grammar strings that contain prospection are not part of conventional literary analysis, and are not at all based on introspection, but they nevertheless do yield to replicable corpus analysis. Louw's corpus-derived subtext (Louw 2013; Louw and Milojkovic 2015, 2016) reveals themes and motifs of short poems (and even students' essays) that are to appear further in the text (see Section 13). Being fully opaque to intuition, this method (reading grammar strings anew with their most frequent lexical variables and not those of the author's choice) is completely automatic and can have the predictive quality that Toolan (2009) does not aim at in his understanding of prospection.

Louw's prospection is work in progress as much as Toolan's is, and both have offered evidence in favour of their model. Still, the predictive quality of Louw's prospection is empirically verified further in the text, whereas Toolan's equating prospection with reader's expectations means, as he himself suggests, that verification rests with studying readers' responses (Toolan 2009: 169). It is also important to note that Toolan's study provides an important link with Louw's view of prospection as potentially encompassing the whole of authorial text. This was not foreseen by Sinclair (2004: 90). Toolan believes

in the potential of prospection to reach across 'very many sentences, to the end of the text' (Toolan 2009: 18).

3.2.10 Philip (2011)

Philip (2011) is the corpus linguist working in the neo-Firthian tradition who is perhaps the most aware of Louw's (as distinct from Sinclair's) particular influence in the area. Creativity, according to Philip (2011: 4-5) hinges on the exploitation of conventional language and is constrained by it. She is a strong advocate of Louw's views, and her contribution is particularly helpful within this discussion as it settles much of the debate mentioned here that oversimplifies Louw's initial paper. The author claims that semantic prosodies are uncontrollable because they are pre-lexical, a pre-verbalised meaning: 'we could risk saying that while semantic associations come from the head, semantic prosody essentially comes from the heart, which is why they are so difficult to manipulate'. She recognises that, unlike connotation and evaluation, SP is only revealed in its context of use. The role of intuition is limited to cases where breaches of SP draw it to our attention.[12] The author's reliance on these premises, as well as on Louw's delexicalisation-relexicalisation continuum (Philip 2011: 120) define her as emerging from the school not only of Sinclair, but also of Louw.

3.2.11 Concluding remarks

All in all, the academic discussion surrounding semantic prosody to date has revolved around several questions: how to name it, how it differs from semantic preference, whether it is a feature of the word or a multi-word unit, and, when it is breached, whether the breach is bound to point to either irony/creative meaning or insincerity, or whether it may be simply neutral. Re-naming it proves challenging because, as pointed out by Morley and Partington (2009), SP is a characteristic of a lexical item in general, but it also has cohesive function in discourse. As for the debate around the distinction between semantic preference and semantic prosody, Partington (2004) is correct in giving semantic preference

[12] Louw (personal communication) also believes that intuition's role is restricted to the choice of the pattern to study. I would go further and suggest that ideally all word combinations should be studied for hidden meanings in authorial text, precisely because our intuitions cannot recognise all such clashes, or other important nuances, such as telling lexico-grammatical collocations.

a lower status than semantic prosody, in which he is seconded by Bednarek (2008), whose analysis confirms the status of semantic preference in relation to semantic prosody defined by Sinclair.

Sinclair was a lexicographer, whose main task was to identify the typical. On the other hand, basing his analyses on authorial texts, Louw was bound to encounter the untypical, simply because any text will differ from the most frequent combinations in the reference corpus. Dwelling on the untypical word combinations, he showed how the typical helped in decoding the untypical. I maintain that it was then that he unwittingly created the impression in linguistic circles that his primary point of departure was the word and that therefore he would not accept Sinclair's idea that a semantic prosody had a scope covering several words. To this overly simplistic view of Louw's notions was added the conviction, expressed in Hunston (2007) and seconded by McEnery and Hardie (2012), that a prosodic clash could explain a readerly reaction, but could not guarantee that the writer must have been either lying or joking. This speaks of the general tendency traceable within corpus linguistics as a discipline to take little notice of the context of situation.[13] A breach of a well-established SP will have its cause. However, this indicates that the apparently 'neutral' cases of prosodic clashes should be researched in much more depth as a separate phenomenon.

In principle, this section has shown that corpus linguistics as a discipline primarily concerned with repeated patterns has difficulty comprehending fully the phenomenon of SP because SP is subject to enormous variation and is difficult to classify. Researching the behaviour of each combination in context combined with cautious generalisation would clearly be a better course of action. It seems that corpus stylistics, and not corpus

[13]Corpus pragmatics caters for this concern: 'Corpus pragmatics is distinct from other fields in corpus linguistics. However, in common with other fields, corpus pragmatics investigates the co-textual patterns of a linguistic item or items, which encompasses lexico-grammatical features such as *collocation or semantic prosody*. However, where corpus pragmatics' "added value" lies is in its insistence that these patterns be considered in light of the context – the situational, interpersonal, and cultural knowledge that interactional participants share. Through an iterative process, corpus pragmatics therefore moves beyond important but surface observations of lexico-grammatical patterns to allow a more nuanced interpretation of these patterns taking into consideration who uses them, where they were used, for what purposes, and how this use has changed over time. In this way, corpus pragmatics has retained in part its original interpretative nature but has endeavored to supply this interpretation with objective supporting evidence' (Clancy and O'Keefe 2015, my italics).

linguistics, is a discipline primarily suited to the exploration of SP, and in particular Contextual Prosodic Theory, because it compares particular authorial usages with the reference corpus norm in *similar contexts of situation*. The next section will review the impact of Louw's views on the discipline of corpus stylistics.

3.3. Louw's impact on the discipline of corpus stylistics

3.3.1 Stylistics and meaning

Simpson (2014: 3-4) defines stylistics as 'a method of textual interpretation in which primacy of place is assigned to language'. Linguistic structure is studied because it reveals the function of the text under consideration and helps the analyst to explain why 'certain types of meaning are possible'. The preferred subject of stylistics is literature, but stylistic analysis ought to be rigorous, retrievable and replicable. Stylistics mostly studies creativity in language, but paying attention to texts in which language rules are 'bent' also helps us better understand the rules themselves.

Thus, the focus of stylistics is textual interpretation through determining the text's linguistic features, which may in addition lead us to draw conclusions that are purely linguistic rather than literary. A similar point is made in Jeffries and McIntyre (2010: 4). Toolan (2018: 60) seconds this view, preferring to focus on 'understanding of how texts work' rather than 'interpretation' in his definition, broadening it to include the communicative dimension. He calls various branches of stylistics 'highly contextualised linguistics'. Talking of 'meanings readers derive from a literary text' and 'projected meanings', he agrees with Simpson's view of stylistics as a study of 'possible meanings' rather than of *the* meaning of a text. To Toolan, the *style* of a particular longer text is its most significant linguistic patterns or breaches of the linguistic norm (Toolan 2018: 61-63).

It can be summarised that both sources define stylistics as the objective and contextualised study of linguistic patterns or breaches thereof in authorial text with a view to 'interpreting' it or understanding its effect on the reader. In contrast to these stylisticians, who cautiously talk of pluralised 'meanings' or 'certain types of meaning', especially from readerly perspectives, Sinclair took the view that 'evidence of shades of meaning and subtleties of expression' could now be obtained by consulting large general corpora. 'Louw

(1993) has given a lead in the literary field' (Sinclair 2004: 117). A stated in Section 2.1, the logic behind Sinclair's optimism drove Louw to establish, relying on the COBUILD Main and Reserve Corpora (approximately 22 million words of running text), the discipline of corpus stylistics, 'at St Hilda's College Oxford in April, 1987 and in the presence of Ron Carter' (Louw, personal communication; see also Louw and Milojkovic 2014: 268).

Currently stylistics is a discipline that encompasses a variety of approaches, such as cognitive stylistics (Stockwell 2002, 2009; Gavins 2013), pragmatics (Chapman and Clark 2014, Clark 2014), critical stylistics (Jeffries 2010), pedagogical stylistics (Fogal 2015, Zyngier and Fialho 2010, Hall 2014).

Studying style, of course, does not come down to merely comparing a text against reference corpora. Reference corpora help to reveal 'latent patterning' (Hunston 2011: 55-56). When it comes to literary texts, conscious effort invested by authors into constructing meaning through *form* can at times be very original. For example, the notion of iconicity is purely stylistic because it emphasises the conscious endeavour invested by the author of a literary text to fit his or her intended meaning into a form that would in turn fit the meaning: 'Analogy or resemblance is invoked: the linguistic form is said to reflect, match, enact, perform, or mime the meaning' (Toolan 2018: 65). I am grateful to prof. Branko Tošović of the University of Gratz for remarking after my presentation on Brodsky's poetry at the Vinogradov Russian Language Institute, Moscow, that one ought to take into account the constraints imposed on authorial choice by the metre and the rhyming pattern. Still, the final product will contain lexico-grammatical combinations that in the reference corpus (if it turns out to contain them) will be found to express certain semantic nuances; poetic corpora and those of the poet under consideration must also be taken into account (as will be shown in this study, see Section 5). Consequently, meaning will be encompassed by the notion of style except in cases where the linguistic form is markedly at odds with the norm as reflected by the corpus (Toolan 2018: 66 calls the latter 'a degree of calculated fit'). This is how Fischer-Starcke puts it:

> Stylistics is the linguistic analysis of literary texts. Corpus linguistics is the electronic analysis of language data. The combination of both disciplines is corpus

stylistics, the linguistic analysis of electronically stored literary texts. Corpus stylistics pursues two goals:

1. to study how meaning is encoded in language and to develop appropriate working techniques to decode those meanings, and
2. to study the literary meanings of texts (Fischer-Starcke 2010: 1).

Fischer-Starcke (2010: 39) quotes Mukherjee (2005: 1184) on style as choice, as deviation from the norm, as recurrence, and as comparison. Clearly, Louw, in his emphasis on checking an author's usage against the reference corpus as representative of the language norm, subscribes first and foremost to the definition of style as deviation (although style as recurrence has also been researched by CPT, see Milojkovic 2011a).

This goes back to the Russian Formalists' notion of literary language as a deviation from the norm, in Shklovsky's (1919) terms 'estrangement' or 'defamiliarisation'. Estrangement (остранение) is a characteristic of poetic language and should yield to scientific analysis. Deviation is central to *foregrounding*, which Dowthwaite (2000: 110) defines as a linguistic element that is salient because it differs from ordinary use. However, despite Mukařovský's (1970: 43, quoted in Siepman 2015: 363) claims that foregrounding is an essential quality of literary language, it is widely accepted in linguistics that literary language is not special (Mahlberg 2015: 347). Louw's focus on exceptions to the norm, however, is not limited to salience, although he does concede that an analyst may intuitively select for study those combinations that have caught his or her eye. Any sequence may conceal deviation from the corpus norm. Also, one of the major goals of this thesis is to illustrate the mechanism of corpus-derived subtext, which is based in syntax and is therefore completely opaque to intuition (Louw in press).

3.3.2 What is corpus stylistics?

Mahlberg (2015: 352) asserts that what separates corpus stylistics from corpus linguistics and computational stylistics (with which it shares its quantitative dimension) is 'its intrinsic explanatory purpose that makes it possible to focus on specific meanings in texts'. To her, corpus stylistics is mainly about literature: 'To delimit the object of study for corpus stylistics it might be more useful to refer to “meanings in literary texts” than

“literary style” (Mahlberg 2015: 347). Biber (2011) distinguishes three major methodological approaches within this field: keyword analysis, identifying typical extended phrases, and collocational analysis. As to the collocational approach (studying the tendency of words to co-occur more frequently than by chance), Biber particularly dwells on semantic prosody (quoting Louw 1993 on 'bent on'), which he finds to be 'very useful for explaining the stylistic effect of particular expressions in a literary work' (ibid.: 18). He notes that 'these studies [e.g. O’Halloran 2007a] have quite different research goals from the keyword investigations surveyed above, focusing on the stylistic effect of a single expression or sentence rather than the overall stylistic characteristic of a text or body of texs’ (ibid.: 19).

Similarly, Simpson distinguishes between two main stylistic applications of corpus techniques:[14] 'One [...] is to set patterns in a literary text against those found in general corpora of the language [...A]ny discussion of *foregrounding* or *deviation* in literature is after all contingent upon, and relative to, some form of quantitative observation about what is 'normal' in language. Another stylistic application is when the corpus itself comprises literary works, perhaps the entire output of a single writer that might run into the millions of words. This facilitates an evidence-based description of the writer's *style* (Simpson 2014: 48-49, my italics).

Speaking of the former kind of corpus methods employed in stylistics, Simpson believes that collocation has 'far-reaching implications for stylistic analysis because to talk of the "meaning" of a piece of language is contingent on how real people actually use it in connected patterns of discourse and in real contexts of use' (ibid.: 48). 'Stylistic applications of semantic prosody are useful because, on the one hand, they support the idea of literature as *difference* and, on the other, help probe the idea of *subtext* in literary composition' (Simpson 2014: 101). The author chooses Louw and Milojkovic (2014) as an example of such a search revealing not only semantic prosody, but also corpus-derived subtext.

As to the latter type of corpus stylistic approaches, it is also employed by CPT, as is exemplified by Milojkovic's (2011a) study on Larkin: the poet's recurrent patterns are

[14] ‘That corpus methods are becoming increasingly relevant is also reflected by their inclusion in the second edition of Simpson’s (2014) influential textbook’ (Stockwell and Mahlberg 2015: 131).

checked in the whole corpus of his poetic works. What differentiates Louw's approach from e.g. that of Mahlberg (2013) is that she compiles reference corpora on the criteria of authorship, genre and time of writing, and that she uses clusters, or lexical bundles - repeated patterns in the given author (Dickens) - which she later checks against reference corpora compiled of 19th century fiction. This is a kind of limited exploitation of semantic prosody, without deploying the possibilities offered by lexico-grammatical collocation as corpus-derived subtext (Wang and Humblé 2017: 553).

Generally speaking, the possibility of using SP and the subtext of grammatical strings in stylistic analysis has not become a central concern of this corpus stylistics. On the contrary, in this discipline emphasis has been rather on repeated patterns retrievable by statistical means. Corpus stylistics has often been understood as corpus techniques normally used within a corpus of (a writer's) works, which can be supplemented by a reference corpus that shares some features with (the writer's) corpus (e.g. Tabata 2002, Hardie 2007, Moon 2007, Culpeper 2009, Toolan 2009, Fischer-Starcke 2010, Siepmann 2015): basically as corpus linguistic strategies applied to the study of literature (such an approach may be combined with cognitive approaches, see e.g. Stockwell and Mahlberg 2015). General reference corpora are sometimes, but not always, taken into account (e.g. Moon 2007). Biber (2011: 17) believes that they may be avoided in order to prevent confusion between register- or genre-related characteristics and the stylistic features that are being analysed. This is a valid concern, but readers decode meanings in texts against the benchmark of the general linguistic norm. Stubbs (2014), describing quantitative methods in stylistics, does mention the use of general corpora, though again not as the researcher's primary concern.

Prior to the publication of Simpson (2014), whose first quoted example of how to apply corpora in stylistics was Louw (2011), Louw's work generally, and semantic prosody in particular (with no mention of corpus-derived subtext) was treated in the literature as one of the many phenomena of interest. For example, unlike Biber (2011: 18), Mahlberg (2013: 19) expresses reservations when it comes to the use of semantic prosody in corpus stylistics: 'Such comparisons can be very insightful for the study of individual items or text extracts. However, there is also the danger that the analysis is only relevant to very isolated

phenomena.' This approach is countered in a recent study, where semantic prosody as such has been recognised as an important analytical tool: 'concepts relating to co-occurrence, including collocation, semantic prosody and semantic preference could be helpful [...] to account for certain patterns of behaviour and particularly in organizing large numbers of examples into data-driven categories' (Jeffries and Walker 2017: 190).

With the publication of Louw and Milojkovic (2014, 2016) the situation seems to have changed, with Louw's understanding of corpus stylistics being recognised as one of its main branches not only in Feng and Humblé (2017). Mahlberg (2015: 353-354) cites Louw (1993) as one of the four main illustrations of the corpus stylistic approach, the other three being Semino and Short (2004), Toolan (2009) and Stubbs (2005). Still, echoing her own remark in the 2013 book, she expresses reservations as to the potential of SP to study whole novels (ibid.). The fact that a consensus has been reached is confirmed by mentions of Louw and Milojkovic (2014) in encyclopedic editions (see Stockwell 2015; Toolan 2015). Still, applications of Louw's CPT, in particular corpus-derived subtext, by authors other than Louw and Milojkovic are only just emerging (e.g. see Castiglione 2019: 308-9, where he searches COCA for quasi-propositional variables in *never to be* *).

Despite the recognition of Louw's theory in the field of corpus stylistics, which he himself established (Louw and Milojkovic 2014: 263), corpus linguistics literature still has not commented on Louw's corpus-derived subtext, explicated in Louw and Milojkovic (2016) and other sources. Semantic prosody, however, seems to have finally received settled judgement in Stefanowitsch (forthcoming), while being referred to by its original name, albeit again with some modifications. Referring to SP's capacity to be evaluative in some cases, and specific in some others, the author refers to them as 'broad' and 'narrow' (Stefanowitsch forthcoming: 231).

3.3.3 Summary

This section has described the development of major views on semantic prosody in corpus linguistics and stylistics, with particular emphasis on the reception of the work of Louw. It has shown that Louw's views found inspiration and support in those of Sinclair, who considered semantic prosody hidden from intuition but an essential element of a

lexical item. Corpus linguistic discussions mainly focused on the difference between semantic prosody and semantic preference (semantic sets that an item might collocate with in the corpus), and on whether an SP was the property of a word or a longer sequence. These questions arose because corpus linguistics as a discipline is interested in phenomena that are typical in the language. Some linguists ascribed to Louw a desire to view SP as the property of a word as opposed to a sequence of text, and may have oversimplified his views as to how to treat exceptions. This was due to corpus linguistics being interested primarily in the structure and distribution of recurrent patterns, and the necessary contradiction that characterises methodologies that would combine statistics with an awareness of the context of situation in each particular case (advocated for by Hunston 2011: 51, who refers to these approaches as 'quantitative' vs. 'qualitative' corpus linguistics). The latter seems to be a contradiction that corpus stylistics can resolve.

4. How ‘shareable’ is textual meaning?

4.1 Is meaning in the text or in the reader?

The previous section discussed the reluctance of stylisticians to talk of 'meaning' in the singular form. Fischer-Starcke summarises linguistic views on meaning:

> 1. meaning is inherent in the language system (Saussure 1916)
> 2. meaning is inherent in a text as a unit (systemic-functional approaches)
> 3. meaning is encoded in recurrent linguistic patterns in a text (corpus linguistics)
> 4. meaning exists in human perception and is projected onto a text (cognitive linguistics) (Fischer-Starcke 2010: 37).

Briefly, these are the views of CPT on these viewpoints. As for Saussure's distinction between *langue* and *parole*, it is only as *parole* that the language system can be studied, via large and representative reference corpora. As to systemic functional grammar, CPT emphasises that a linguist's interpretation of a text, without taking reference corpora into account, remains intuitive. Louw is also critical of Halliday's locating collocation within syntax and colligation, contrary to Firth's (1957: 196) understanding of collocation as abstracted at the level of syntax (see also Louw and Milojkovic 2016: 40-43, and Feng and Humblé 2017: 553-554 for a critical response to Louw's view). As for recurrence of patterns, it is indeed part of CPT's methodology: recurrence in the reference corpus is a basis of generalisations, but these are then employed to clarify semantic nuances in authorial text, which is the starting point and goal of a CPT-driven approach.

As to the fourth outlook as described by Fischer-Starcke, it is a premise of Contextual Prosodic Theory that meaning resides in the text. CPT is skeptical of accounts of what goes on in the human mind because the human mind, unlike the corpus, is inaccessible. This by no means underestimates research conducted in the area of cognition, but is stated in order to delineate the remit of CPT: meaning that can be inferred on the basis of text. However, we cannot completely ignore the fact that individuals process text differently. This section will describe how Louw and Milojkovic (2015: 536-541) deal with this issue.

4.2 Readers' responses to foregrounding: experimental studies

Stylistics, with its emphasis on rigorous precision, is wary of statements that a certain text 'means' (see Section 3.3.1). Indeed, given that even native speakers of the same language vary significantly in brain structure, character traits, personal histories, education levels and backgrounds, talents, and many more aspects crucial to the process of text interpretation, it would be reasonable to doubt that very much could be shared by different people reading the same text, and to assume that reactions and opinions regarding texts may differ substantially. Village (2007: 105-108) recounts the development of the theory of psychological type, which is relevant to the way in which people relate to the world, perceive information and make decisions, and which was first put forward by Carl Jung as far back as the 1920s. While Village discusses this in relation to hermeneutics practiced by laymen, this certainly has a bearing on any text interpretation. According to Village, after reviewing the writings of poets and philosophers, Jung came to distinguish between introvert and extrovert attitudes to the processes of perceiving (i.e. acquiring information) and judging (i.e. evaluating information). Perceiving has to do with gathering information (sensing) and combining it into novel combinations (intuition), and judging may operate 'from the heart' (through 'rational feeling' and values), and 'from the head' (through logic and 'rational thinking'). Subsequent research into psychological types showed through observation and measurements that people express preference for either extroversion or introversion, either sensing or intuition, either feeling or thinking, and, adding the fourth dimension, either perceiving or judging when relating to the outer world. Thus, sixteen psychological types were revealed.

Since these types approach text interpretation differently, how 'shareable' are readers' reactions to texts? Encouraging findings have been gained through empirical stylistic research into reader responses. Miall and Kuiken (1994, 2001, 2002) investigate how readers process foregrounding (defamiliarisation), and propose the notion of refamiliarisation – the reader processes the foregrounded passage by comparing it with other passages from the same text, or, for example, recalling personal experiences. The passages in question are deemed foregrounded without recourse to corpora in these studies,

and the authors add that 'such reconsideration of the text surrounding foregrounded features will be guided by the feelings that have been evoked in response to those features' (Miall and Kuiken 1994). In later research (2001, 2002), the authors call this process're-contextualisation', connecting the aesthetic feelings aroused by foregrounding to a feeling of uncertainty that forces the reader to dwell on it and resolve the tension. Fialho (2007) recounts the research described above and proceeds to set up a further empirical experiment intended to clarify how readers process foregrounding. Her conclusions are supportive of preceding research. In her experiment, three different groups of subjects pointed to the same passages, taken from two short stories on the theme of love, as foregrounded. These three groups were six independent evaluators, fifteen advanced literature students and fifteen students of engineering. Although the foregrounded passages were first chosen by the six evaluators without recourse to corpora, '[t]he segments with the highest indexes of foregrounding were the ones that elicited comments more often. The segments with the lowest indexes of foregrounding were the ones less frequently chosen for commentary, by both [investigated] groups' (Fialho 2007: 119). The other characteristic shared by the advanced literature students group and the engineering students group was that, although the latter was shown to possess far fewer interpretation strategies, the majority in both groups focused on the text itself (or did not position themselves at all) and few subjects made the transition from textual engagement to extra-textual reflection. Finally, when it comes to feelings expressed (termed 'affect' in the study), the majority's reaction was either indifference or uncertainty. Overall, despite possible misgivings related to diverse reactions coming from diverse readers with diverse backgrounds, it seems that neither psychological types nor education and backgrounds have a defining influence on readers' perception of literary texts, and, more significantly, on what they consider foregrounded.

4.3 Are meanings dependent on patterns: a linguistic perspective

Closer to home, in linguistics the idea of the inter-relatedness between patterns and meaning is Jakobson's (1958: 20). According to him, '[t]he poetic resources concealed in the morphological and syntactic structure of language – briefly, the poetry of grammar and its literary product, the grammar of poetry – have been seldom known to critics and mostly

disregarded by linguists but skilfully mastered by creative writers.' This idea is taken up by Halliday (1971), who illustrates how purely linguistic features are used in characterization (discussed in Bettina Fischer-Starcke, 2010: 45). Fischer-Starcke goes on to suggest that this analysis 'assumes that linguistic features and patterns in a text have meanings and evoke meanings for the reader'. Halliday's analysis of William Golding's *The Inheritors* is not aided by computers: 'He demonstrates that the description of Lok's tribe frequently uses simple past tense forms, has a preference for non-human subjects and that transitive verbs are almost completely absent. The linguistic patterns in the description of this people create the impression for the reader that the tribe is both inefficient and helpless in its actions.' The hypothesis is borne out by the plot: 'And in fact, the tribe is attacked and defeated by another tribe in the story' (ibid.).

On the other hand, the discipline of corpus stylistics, or at least its branch started by Louw in 1987, offers tangible proof of native speakers' reactions to texts and deviations from the language norm in them, whether the deviations are noticed on the conscious level and are such that make the reader stop and process them, or whether they influence the reader subliminally rather than consciously. It is the latter contribution - revealing deviations from the language norm that have taken place but are not consciously processed - that singles out CPT as particularly relevant to stylistics as a branch of linguistics concerned with textual interpretation. The empirical proof that a deviation has taken place lies in the comparison between a language pattern used in a text under discussion and the way in which the same language pattern is used in the general reference corpus. O'Halloran (2007b, discussed in Mahlberg 2013) deals with the subject of readers' reactions along similar lines: a reference corpus will predict that '[i]f the corpus comparison points to deviations from typical language schemata, this can be taken as evidence that a tension between the actual text and typical world schemata will occur in the process of reading. As the corpus provides evidence for language patterns that are associated with shared world schemata, it is further possible to assume that the tension will be felt not only by one reader but by readers in general' (Mahlberg, 2013: 38). Contextual Prosodic Theory (CPT, see Louw 2000) subsists in contexts of situation and states of affairs in both authorial text and the corresponding patterns in the reference corpora, and these seem to take account of the

abstract ‘typical world schemata’ employed by O’Halloran (2007b). It must be stressed at this point that this thesis by no means denies the usefulness of cognitive studies and psycholinguistics. What it hopes to bring into question is the theoretical assumption that the experience of an individual has primacy over the semantic auras of lexis and grammar that are empirically verifiable. As regards language schemata, text schemata, and world schemata, one may make inferences regarding these on the basis of reference corpora if one so wishes, but here a similar argument applies: the actual reference corpus contexts are more immediately connected to a particular text sharing the same pattern with them than abstract concepts such as schemata. As Frank Palmer used to put it, where data is plentiful, concepts are superfluous (Bill Louw, personal communication).

4.4 Research of semantic prosody in experimental psychology

That meaning is shareable through semantic prosodies has recently been verified by means of experimental psychology. Hauser and Schwartz (2016) describe a comprehensive study they conducted to discover whether semantic prosody can affect inferences drawn from an utterance. The underlying assumption was that a described event would be evaluated more negatively or positively if the description included a term with a negative semantic prosody, and vice versa. The words in question were 'attain', 'lack', 'restore', 'lend', 'emphasise', taken to have a positive SP, and 'cause', 'encounger', 'commit', 'arouse' having a negative SP. According to the authors, these words were chosen because their attitudinal component - or valence[15] - was not part of their dictionary definitions, which they had checked. They then identified neutral pairs for these: 'get', 'not have', 'bring back', 'loan', 'stress' on the one hand, and 'produce', 'happen upon', 'engage in', and 'evoke' on the other. The study showed that when it came to conscious awareness of a semantic prosody, it was not exhibited even in the case of 'cause'. But when the subjects were asked to evaluate diagnoses, the phrasings of which contained semantic prosodies, they did evaluate

[15] ‘Valence’ in psychology means the intrinsic attractiveness or averseness of an object or a situation (Frijda 1986: 207).

more negatively those phrasings which contained negative SPs. Also, for example, participants evaluated more positively the context in which 'the senator produced' something, than the one in which 'the senator caused' something. The methods of the study varied and the numbers of participants ranged between 40 and 1177. The study, being far more comprehensive than can be described here, is a most valuable contribution to research on SP: 'Our results also have important implications for the field of semantic prosody. Although semantic prosody has been widely documented in natural language (Louw 1993; Stubbs, 1995; Partington, 2004), some noted a lack of evidence that the semantic prosody of a word predicts a reader's *interpretation* of valence in novel contexts (cf. Hunston, 2007; Stewart, 2010; Whitsitt, 2005). Our experimental results show that semantic prosody does indeed exert a causal influence on readers' inferences about valence, which are reflected in explicit valence judgments and related inferences. Moreover, native speakers completed sentence fragments in a manner that makes them consistent with valence patterns of co-occurrence (Study 2 and Study 4). Thus, evidence for a causal influence of semantic prosody is no longer lacking' (Houser and Schwartz 2016: 914). A similar study on different SPs was conducted by these authors in 2018, which supported the findings obtained in 2016: We conclude that semantically prosodic descriptors that lack a clear positive or negative meaning still lead people to infer the valence of what is to come, which colors the impressions they form of others (Hauser and Schwartz 2018: 11).

4.5 The interaction of semantic auras within a text

The newest development in CPT, namely logical semantic prosody – subtext, focuses on grammar (logic) rather than vocabulary (metaphysics). Logical semantic prosody – subtext – introduces a new level of analysis, as it focuses on grammatical strings instead of lexis. Grammar strings, whose most frequent lexical collocates – their quasi-propositional variables, or QPVs (Louw, 2010a, 2010b) – determine their corpus-derived subtexts. Both Hoey (2005) and O'Halloran (2007a, 2007b) deal primarily, though not exclusively, with lexis and semantics. As opposed to lexis, grammatical strings have semantic auras which are completely opaque to intuition, and are not easily defended either

by the hypothesis of psychological priming, or by typical expectations seen as schemata – except in a very general sense. And yet, grammar strings have consistent subtexts, which interrelate within, say, a poem – which is successfully illustrated in Louw and Milojkovic (2014) and will be at length defended in this thesis in its subsequent sections (see e.g. Section 7). Moreover, the most recent development in CPT takes subtext further by suggesting that the first lines of poems – and not only poems – prospect (Louw 2013) developments in texts. This empirical fact, as well as the general interconnectedness of subtexts within a poem, together with subtext's intuitive opacity, rather suggests that the fabric of the language underlies language use, whether receptive or productive, and that the mechanisms of that use are mostly subliminal and only recoverable through corpora. These all-pervasive mechanisms can be used to discover additional meaning in any text, as any text is bound to differ from all other texts containing the same linguistic patterns. If foregrounding, which produces tension in the reader, may be viewed as forms of deviation so significant that it is spotted with the naked eye, this does not mean that smaller and less obvious deviations may not result in tensions. Still, foregrounding has been shown empirically to be spotted and processed similarly by people with different backgrounds and necessarily different personal psychological primings, therefore there is every reason to assume that smaller deviations must be processed at some level, or language experience would not display the consistency it does. As previously mentioned, the consistency of semantic auras of grammar will also be illustrated and defended in the subsequent sections of this dissertation.

To illustrate this consistency, let us look at a poem by Philip Larkin, and focus on the stretch of text highlighted in bold:

When first we met, and touching showed
How well we knew the early moves
Behind the moonlight and the frost
The excitement and the gratitude
There stood how much our meeting owed
To other meetings, other loves.

The decades of a different life
That opened past your inch-close eyes
Belonged to others, lavished, lost;
Nor could I hold you hard enough
To call my years of hunger-strife
Back for your mouth to colonise.

Admitted; and the pain is real.
But when did love not try to change **Is this an optimistic reference to love?**
The world back to itself – no cost,
No past, no people else at all –
Only what meeting made us feel,
So new, and gentle-sharp, and strange?

These are some of the contexts from the reference corpus of the 1995 edition of the Times newspaper:

1 Banks, insurance brokers and estate agents sell their products and there's nothing wrong with that. **But when did a car salesman ever tell you that you would be better off walking or taking a bus?**

2 A politically imperilled Government will probably still opt to cut taxes instead. This may make little economic sense, **but when did economics really come into the equation so close to a general election?**

3 TOMMY BOY, 97 mins, PG
After Dumb and Dumber, we now have Dumbest to date. Starring Chris Farley, yet another dubious Saturday Night Live Graduate, this is not so much a comedy of errors as an error of comedy as our hero takes over the family car-brake business when his father (the much-abused Brian Dennehy) dies from over-exertion caused by marrying Bo Derek. Dan Aykroyd and Rob Lowe also participate, **but when did either last make a prudential career move?**

6 The green devotees will doubtless tune in to Witness: Beyond the Rainbow (C4, Wednesday, 9 pm), in which the daughter of a photographer killed in the sinking of the Rainbow Worrior embarks on a quest to find out more. This approach to documentary-making virtually ensures partiality, and has become a cliché. **But when did that ever deter anybody**?

And this is the whole concordance from which the lines were taken (Milojkovic, 2012):

MicroConcord search SW: but when did

1 there's nothing wrong with that. But when did a car salesman ever tell you that
2 may make little economic sense, but when did economics really come into the eq
3 may make little economic sense, but when did economics really come into the eq
4 d and Rob Lowe also participate, but when did either last make a prudential car
5 ies in both manager and country. But when did England last have success or a pu
6 be the logical time to bow out. But when did football, life, logic, Charlton o
7 tiality, and has become a cliche. But when did that deter anybody? </Group>
8 abs. The rot set in after that. But when did the present system start, and why
9 ompany, making £2billion a year. But when did you last hear critics sounding of
10 hormone is, of course, a cop-out,but when did you last hear of a netball crowd

Figure 4.1. The Times 1995 concordance of 'but when did'

With the exception of one line containing a question, the rest are all rhetorical questions, and they all refer to previous experience in order to support criticism, the context being strictly negative. But will the subtext of the first line of the poem substantiate this lack of optimism? Does the first line prospect the negativity suggested by corpus findings? Below are the four contexts of ‘when first we *’ obtained from the BNC:

1 But, I, what I can't understand is, why all of a sudden does he want her back and see the children, well for over a year he kept her away? Told the children lies. (SP:PS1HJ)Unless he doesn't know she living with someone.(SP:PS1HH) You know that saying Jean? (SP:PS1HJ)(unclear)(SP:PS1HH) No. Oh what a tangled web we weave (pause)**when first we** practice to deceive. (SP:PS1HJ)

2 It was not for some years after we had separated that I had to face the fact that the deepest part of one's being is always inaccessible to another, and that only when we are lost in passion, eager above all else for our own gratification -- yet unconscious that this is the case -- that we ignore the gulf, forget that we are strangers. That part of me which, **when first we** met, I should willingly have made over to Jean-Claude -- and which he seemed to make no attempt to acquire -- was, of course, no more available to him than the source of his being was to me.

3 What could I have to gain, other than making the best of a bad bargain?'' I know, I know,' he said soothingly.' If I have a motive,' she pursued, her belligerence lessening,' it is no different from that which led you to aid me **when first we** met. The poor woman was left in the wretchedest discomfort. If I feel compassion for her, it is not to be wondered at.' She found that tears were trickling from her eyes, and dashed them away with an impatient hand.

4 The nations of the world have become less and less disposed to help each other. Certainly the situation will not be improved by a price war in oil. Nigeria, for example, produces oil which is similar to ours. Nigerians accuse Britain of leading a cut-price war against them. That was their view **when first we** reduced our North Sea oil prices, and now they think it again. So what? Just this: Nigeria is our major trading partner in Africa. Our trade with Nigeria is greater than that with South Africa. Thus an enormous amount is at stake in our relationship with Nigeria. We should be talking with them and going out of our way to assure them that we do not wish to engage in a price war.

Although the contexts depict diverse (at first sight) contexts of situation, the negativity is obviously present. Three contexts of situation are centred on participants of a love relationship. In context 1 separation and deceit are discussed. In context 2, separation gets a mention first, but the stretch of text in question refers to the notion that, in a love relationship, sharing oneself completely with the loved one is an illusion. Context 3 is a lovers' quarrel. Unlike these, context 4 does not refer to man-woman love, but to international trade. Still, the situation does involve two parties, one suspecting the other of unfair competition. Contexts 2 and 3, apart from being focused on man-woman love, also feature the lexical (quasi-propositional) variable 'met', the same as in Larkin's text. While context 3 is focused on a misunderstanding, which the reader feels is about to be patched up, Context 2 speaks about relationships more generally and more pessimistically: no matter how much one wants it, the complete sharing of oneself with one's romantic partner is an illusion. This links in with the sentiment expressed by Larkin's poem: one more attempt, one more chance at hope. Also, the criticism and anger so overwhelmingly expressed by the Times concordance of 'but when did' may be seen in the contexts 1, 3 and 4 (in context 4 the key collocate is 'accused', in the same line as and preceding 'when first we reduced'). All in all, the subtext of the first line can be linked with that of the highlighted text in the final stanza, even though intuitively this cannot be guessed at. As for the highlighted text and the question pertaining to it, sixteen native speakers out of the seventeen consulted replied that this reference to love was optimistic. This example shows that the fabric of grammatical strings underlying language is consistent, although opaque to intuition. It is not to be wondered at that the native speakers opted for a conclusion refuted by corpus findings. Meaning residing in subtext is only part of the pattern. There are other context clues. As an English native speaker pointed out:

> The use of 'but' at the beginning suggests that the section is intended to stand in contrast to what is 'admitted'. Of course, the use of 'try' in the section accepts the possibility that love may not succeed in its attempt to change the world back into itself, and the use of 'back' implies that the world was once love, but has changed, so the optimism is tinged with a sense of loss and doubt. But the assertion carried in

the way the question is phrased - that love always makes the attempt - suggests to me a hope that will never allow experience to triumph over it.

These context clues may lead to the assertion that Larkin himself intended for the text in question to convey optimism rather than not, and that the resulting aura of loss and doubt is as subliminal in the author's use as it is in the reader's perception. The pessimistic aura surrounding 'when first we' seems to predict this loss and doubt hinted at in the last stanza. The consistency of findings seems to indicate that areas which our intuition cannot see lie at the heart of our reactions to texts in the sense that they are the most shareable. The paradox is that the sharing, of which we must be convinced on the basis of the frequency of patterns, is entirely subconscious. This idea Louw calls 'corpus-derived empathy in stylistics' (Louw and Milojkovic 2015: 535). As Fischer-Starcke puts it when discussing corpus stylistics as a discipline, '[y]et, it is precisely this loss of individuality, that is, a reader's personal textual competence and experiences, that corpus stylistics aims for in the generation of the data that is analysed, as this is what contributes to the intersubjectivity of an analysis. The generation of frequency data as a basis of the analysis of literary meanings is as much stripped off an analyst's individual choices and perceptions as possible' (Fischer-Starcke, 2010: 7).

4.6 General semantics or psychological association?

The phenomenon of semantic auras interconnecting and interacting in texts will be studied in detail in subsequent sections. This section has aimed at clarifying what can be considered explanatory, and what predictive, when it comes to the power of semantic prosody and corpus-derived subtext to construct meaning. What goes on in the readers' heads CPT cannot predict, as much of it lies in the domain of psychological association. However, psychological associations need to be distinguished from general semantics. The reference corpus, under the condition that it is accepted as the language norm (in the hope that reference corpora will grow more and more in the future and thus equip the researcher with a more precise instrument of comparison) serves as an objective, and not intuitive, benchmark of how collocations are used in context. This means that Louw's method of analysis provides a sharper understanding of language norms at all levels (collocation,

context of situation, context of culture) for particular word combinations. The inferences made as a result of analysing authorial text against the background of reliable reference corpora have to do with general semantics, as opposed to personal associations. This study in its subsequent sections will give examples of semantic auras consistently supporting one another within texts. At the moment, is it crucial to state that Louw's approach does not allow for semantic nuances of a particular text to be understood merely as arbitrary meaning or one of the possible meanings. If they are questioned, it is in the same way that dictionary definitions should be questioned. 'Whether or not you agree with their dismissal of attempts to describe the conceptual underpinnings of discourse and discourse processing, the authors' [of Louw and Milojkovic (2016)] dedication to data and the scientific method is laudable' (Lugea 2017: 353).The same author gets carried away when she states: 'For Louw and Milojkovic, the text's artfulness is *measurable only in relation to observable data, through corpora or empiricism*' (ibid., my italics). This is a contradiction in terms: if 'artfulness' is at all measurable, the measurement must be based on objective criteria. Psychological associations of various readers, together with their accumulated experience of the language being sufficient or otherwise for the processing of a particular text, do not fall within the remit of CPT. As Teubert (2005: 2-3) puts it: 'The focus of corpus linguistics is on meaning [...] It is not concerned with the psychological aspects of language. It claims no privileged knowledge of the workings of the mind or of an innate language faculty'. Perversely, not being able to predict the working of a particular mind upon encountering a text, procedures deployed by CPT shed light on the full shareable meaning potential of authorial text, including nuances not accessible to human intuition when unequipped by computers.

5 Transferability of Contextual Prosodic Theory to Russian

5.1 Introduction

This section (based on Louw and Milojkovic 2016: 195-240) will focus on the practical implications of Contextual Prosodic Theory for Slavic languages, namely, Russian. The very application to languages other than English points to the theory's transferability and universality. To this end, it will provide practical examples to discuss already familiar notions, like collocation and co-selection leading to states of affairs, semantic prosodies and fractured contexts of situation, and subtext. It will also dwell on relexicalisation and its significance for literary devices. All these notions will be looked at separately at first. However, as textual analysis at its best does not subsist fragmentarily, so all these aspects will be combined to reveal the logical construction of literary worlds, as well as a hitherto unexplored idea in corpus stylistics: authorial intention, clarifiable by lexical collocation and subtext.

5.1.1 Semantic prosodies in Russian

In order to apply Contextual Prosodic Theory to Russian, a useful first step would be to ascertain if there are Russian equivalents to the most famous English semantic prosodies. Let us take as an example the verb 'cause'. According to Stubbs (1995: 24), 'nearly 80% of occurrences have clearly negative collocates, usually within a span of 3:3. Conversely, a very small number of occurrences have positive collocates. The distribution is: negative 80%, neutral 18%, positive 2%'. Arguably, the 2% needs further subdivision in the light of Louw's (1993; 2000) hypothesis that exceptions to the prosody will be comprised of one insincerity to every 2 ironies.

The findings yielded by the Russian National Corpus[16] (henceforth RNC) suggest that the corresponding Russian verb *вызвать* ('cause', perfective aspect) behaves in a similar way. In the search involving all forms of the verb, in the first 103 random lines (those containing the verb in the meaning of 'summon' were excluded), 67 contained negative collocates, 13 contained positive ones, and in 18 lines collocates were neutral. In the positive instances in 6 cases out of 13 the direct object of the verb was *интерес*

[16] All the searches in this section were made in 2011, unless stated otherwise.

(‘interest’), described as *большой, живейший, неожиданный* etc. (‘great’, ‘liveliest’, ‘unexpected’ etc.). The verb *вызывать* (imperfective aspect, the same meaning) behaved even more similarly to its English equivalent: out of 107 lines 80 contained nouns of negative meaning as objects of the verb, 14 contexts were neutral, and 13 lines contained positive collocates (again, *интерес* (‘interest’) appeared in 6 lines out of 14). Other collocates functioning as direct objects of the verb, whether in the perfective or imperfective aspect, were *восторг, одобрение, восхищение, внимание, резонанс* (‘enthusiasm’, ‘approval’, ‘delight’, ‘attention’, ‘response’). All positive usages of the verbs appeared in the language of journalism. The several cases where a positive collocate appeared with a negative form of the verb, e.g. *‘большой радости не вызвало’* (‘did not cause great joy’), *‘особого энтузиазма не вызвало’* (‘did not cause any special enthusiasm’), *‘не вызвало воодушевления’* (‘was not greeted with delight’) were ignored. However, their occurrence suggests that we are dealing with Louw’s category of ironic prosodic clashes.

As the verb *вызвать* (‘cause’, perfective aspect) was used positively only in the language of newspapers, a subcorpus of fiction was selected to see whether the usage was limited to journalistic clichés. This was not the case. In 115 random concordance lines 74 contained the verb used negatively. 7 lines more included positive collocates but the author’s overall tone was negative, e.g. *‘даже старая грымза вызвала в нем тень нежности’* (‘even the old hag invoked in him a shadow of tenderness’) or ironic, e.g. *‘полет любви к нему, вызванный заплатой на курточке’ (‘*a wave of love for him, caused by the patch on the jacket’). 5 lines out of the remaining 14 referred to a number of people, i.e. audiences: the collocates were ‘*хохот зала’* (‘a burst of laughter in the auditorium’), *‘восхищение итальянской публики’* (‘the delight of the Italian audience’) etc. Three lines, where the verb was negated but followed by a positive collocate, were ignored. Six more lines were difficult to classify as they contained expressions *вызвать сочувствие* (‘cause sympathy’) and *вызвать жалость* (‘cause pity’) that seem to be fixed. In the given concordance in all 6 cases the author’s tone was tinged with disapproval caused by the inappropriateness of the sympathy or pity.

The *Times* reference corpus seemed to conform to this tendency. In its 44.5 million words 7 lines containing the collocation ‘cause interest’ were found. The collocates of ‘interest’ were ‘much’ (twice), ‘more… than usual’, ‘the explosion of’, ‘a great deal of, ‘some ‘and ‘widespread’. Thus, though few, the examples show that ‘interest’ collocates with ‘cause’ mainly when its considerable degree is emphasised. The situation in Russian is similar. There were no collocations with ‘attention’, unlike in Russian, a single example contained ‘admiration’, 3 citations contained ‘delight’ and one more featured ‘an outburst of enthusiasm’. No explicit ironies were found.

Another example of equivalence can be found in *случиться* (perfective) and *случаться* (imperfective). These are Russian equivalents of the English verb ‘happen’, which was found to carry negative prosody (Sinclair 1991: 112). Since the concordances offered by the RNC often referred to events described by whole clauses rather than a direct object, e.g. *случилось, что...*(‘it so happened that...’), or *это случалось нечасто* (‘it happened infrequently’), where the pronoun referred to a clause, whole contexts and not only collocates were taken into consideration. The verb *случиться* had a negative meaning in 60% of the cases, positive in 24% and neutral in 16%. *Случаться* was negative in 64 lines out of 104, positive in 11 lines. The remaining lines were particularly difficult to classify, but they were not expressly negative. The negative prosody of only about 60% may suggest that we are dealing with a transitional phase in the diachronic development of the language. Even in English ‘miracles happen’. *The Oxford Collocations Dictionary* (2003) gives ‘happen’ as the first collocate of ‘miracle’.

Let us consider a third example. Louw (2000) proves that the prosody of ‘build up’ is negative if the verb is intransitive and positive if transitive. There may be a simple explanation. What humans ‘build up’ must be useful and constructive. What ‘builds up’ by itself is usually destructive: decay, dirt, ruin, other forms of decline and degeneration; or, in the sphere of emotions, in English it is anxiety and tension. In Russian, among the several possible translations, I chose *накапливаться* (intransitive, imperfective, reflexive) and *накапливать* (transitive, imperfective). The imperfective aspect was chosen since the verb ‘build up’ is more frequently used continuously (in the *Times* corpus the findings were 12 instances to 5).

Predictably and consistently with previous findings, *накапливаться* was used negatively in 64 lines out of 105. It was used positively in 21 lines (the most frequent collocates were *опыт* ('experience'), *данные* ('data'), *средства* ('means')) and neutrally in 17 lines. Again we find a more fully negative English prosody to be around 60% negative in the Russian counterpart. What about the positive prosody of the transitive verb 'build up' and the Russian *накапливать*? Predictably again, 66 lines out of 100 were positive contexts, 10 were negative and 24 neutral.

Table 5.1 The negative and positive semantic prosodies in Russian

verb	negative	positive	neutral
вызвать	65%	12,6%	17,77%
вызывать	74,77%	12,15%	13.01%
вызвать (subcorpus of fiction)	64,34%	6,08%	unclassified
случиться	60%	24%	16%
случаться	61,5%	10,6%	unclassified
накапливаться	60,95%	20%	16,19%
накапливать	10%	66%	24%

These semantic prosodies are obviously primal, as they are consistent in both languages (and also in Serbian, as will be shown). Why then are they less dominant in Russian? Among possibly numerous reasons, a few are immediately obvious. In the case of the negative prosody, usually the verb is either found in negative contexts or it takes several fixed positive collocates, e.g. *вызвать интерес* ('cause interest'), *вызвать восхищение*, ('cause delight'); in the context of this verb it was noted that these fixed collocations have become ironic: e.g. *восторга не вызвало* ('was not greeted with delight'). If this is indeed a diachronic phase, it would seem as if the positive collocates became fixed before the prosody developed more fully. The same with other negative prosodies: in Russian miracles happen (*случаются чудеса)*, while experience, data, means build up (*опыт, данные, средства накапливаются*). The other reason is the grammatical differences in the two languages. While the intransitive 'build up' cannot by definition be used in the passive voice, the Russian *накапливаться* (literally 'accumulate itself') may have a human agent.

For example, *военными накапливаются средства* (‘means are accumulated by the military’, literally ‘by the military accumulate themselves means’) differs grammatically from *в почве накапливаются тяжелые металлы* (‘heavy metals accumulate in the soil’, literally ‘in the soil accumulate themselves heavy metals’). The choice of the Russian equivalent may also be an issue – theoretically, it is not certain that the same pattern may be found in other Russian equivalents of ‘build up’, i.e. *создаваться.*

These 3 examples have been studied in detail so as to illustrate the existence of semantic prosodies in Russian comparable to that of English. This shows that, as things stand, the application of CPT to the corpus stylistic study of Russian texts is justified.

5.1.2 Semantic prosodies in Serbian

The transferability and universality of CPT seem to be supported by semantic prosodies in the case of Serbian, another Slavic language, where they function similarly to those in English and Russian. The Serbian verb *izazvati* (‘to cause’, perfective aspect) behaved similarly to its Russian counterpart *вызвать* (‘cause’). The 100 random lines yielded by the Corpus of Contemporary Serbian were all written by Serbian writers, and the majority of texts were literary (it is important to note that in this random sample there were no contexts written by journalists). Contexts in which the verb was used in its other meaning – ‘to challenge’ – were skipped. I included all the verb forms into the searchline, but not aorist forms and imperfective tense forms, and not future forms. The most reliable finding was the count of cases where *izazvati* was followed or preceded by negative collocates, which were sufficient to pronounce that the verb was used negatively – there were 74 such lines. In 16 lines it was not clear from the context if the evaluative meaning was positive or negative, or both were included. In 10 lines the collocates were positive. One line, where the verb was negated but followed by a positive collocate, was ignored.

The positive collocates were similar to Russian – for example, *veliku radost* (‘great joy’*), gromki smeh* (‘loud laughter’), *smelu nadu* (‘daring hope’). The fact that there were no journalists among the authors may account for a greater diversity of contexts and collocates, a small number of repeated ones (only *radost* (‘joy’) was repeated once, and ‘curiosity’ appeared as *radoznalost* and *ljubopitstvo*), and subtler auras of meaning

surrounding the verb. On the whole, it seems that, when positively used, this verb creates strong emotional reactions, especially in a group of people – this finding, after all, corresponds to what was said of the Russian equivalent verb *вызвать* ('cause'). However, no explicit ironies were found, and there were 5 firmly positive uses of the verb that involved individuals and not groups. *Izazivati*, the Serbian counterpart of 'to cause' in the imperfective aspect (not to be confused with the imperfective tense in the Serbian language, which is an archaic past tense), was analysed under the same conditions as its perfective aspect counterpart, both regarding the verb forms and the exclusion of contexts in which the verb means 'to challenge' and not 'to cause'. It was followed by a negative collocate in 68 contexts out of 100, and by positive in 18 contexts. There were 14 contexts where the verb was followed by both positive and negative collocates (5 cases), by neutral ones (in 7 cases), or the evaluative aspect was not clear from context. In the positive contexts, only one use was obviously ironic; in only 4 cases out of 18 the emotion/reaction was caused in an individual.

A Serbian alternative of 'happen' in the perfective aspect, *desiti se* seemed to correspond to its English counterpart. It was shown to prospect a negative event in 64 lines out of 100, a positive one in 5, and a neutral one in 18.[17] Among the positive ones, there was 1 co-occurrence with *čudo* (miracle). There were 18 unclassified contexts 7 of which referred to a sudden and strange event, mostly bringing change. In the negative contexts, the majority were easy to classify because of expressly negative collocates, e.g. *zločin* ('crime'), *tuča* ('fight'), *nesreća* ('accident'). This differs from the case in point in Russian, where there was a greater need to read whole clauses to ascertain the nature of the prosody. It must be noted that both languages contain other verbs with the same meaning, which may in theory show different behaviour if analysed.

Table 5.2 Semantic prosodies in Serbian

verb	negative	positive	neutral or unclear
izazvati	74%	10%	16%

[17] The same research conditions applied here as with the previous two Serbian verbs - the same verb forms were analysed and contexts where the verb meant *zadesiti se* ('happen to be') were considered neutral.

izazivati	68%	18%	14%
desiti se	64%	5%	36%

The comparison between the Russian and Serbian equivalents of 'happen' analysed here may become significant when we discuss their positive prospection. The number of negative cases is similar, 60% in Russian and 64% in Serbian, but there are almost 5 times more positive usages of the Russian verb. This may be accounted for by the fact that the pages of the Contemporary Corpus of Serbian which appeared in response to my search contained a greater number of contexts from novels by authors such as Andric and Crnjanski than you would ordinarily expect in a balanced reference corpus. These novelists trace the fate of the Serbian nation through their protagonists' stories. This might have influenced the findings, as the historical circumstances described by the greatest Serbian novelists are certainly those of strife and suffering. Alternatively, this may mean that the verb almost never prospects positive events in Serbian in general. This may have to be researched more thoroughly.

On the whole, the best known English semantic prosodies seem to correspond to their counterparts in Russian and Serbian. This fact alone is the first step in proving that CPT is transferable to Slavic languages, as well as to other languages where the existence of SP has been noted. The differences in collocational behaviour, studied in this section, were to be expected. More crucially, the very existence of the mechanism of SP in Russian and Serbian supports the application of collocation as instrumentation for meaning in these languages. Although the subsequent sections will use English examples to prove a point, their primary focus will be on Russian, in particular on the texts of arguably the greatest Russian poet, Alexander Pushkin.

5.1.3 States of affairs, events, transitions

How, then, do we stylistically exploit prosodic clashes in Serbian or Russian? Apparently we need to narrow down the range of what we consider unarguably definite prosody. For instance, if the subject of *накапливаться* ('build up') has a positive meaning, but is not in the semantic fields of data, money or experience (practically the only positive

collocates of the verb), we may start wondering if the mismatch is not accidental. This is how we exploit the notion of 'similar events' (Louw, 2010a). Clearly, events make stylistic analysis of every transition an individual experience, since every individual context of situation will contain unique features.

The difference between states of affairs, events, and transitions needs clarification at this point. ***States of affairs*** are chunked by several co-occurring collocates. Significantly, collocation is taken as abstracted at the level of syntax, provided the collocates are within the same context of situation. ***Events*** in the studied text and similar events in the reference corpora are particular instantiations of states of affairs. Each contextualisation of a state of affairs is viewed as an event. Thus, events are also chunked by collocates used to create them.

A ***transition*** in a text is the particular combination of words which has attracted the stylistician's eye as worthy of investigation. This combination will be interesting because it will contain an event in progress – in fact, the progress of the event will hinge on it. If this combination of words is not unusual in the reference corpus, the event will develop predictably. If not – and that is perhaps more probable because there must be a reason why a particular combination of words has caught the analyst's attention – then it will contain a fractured (deviant) semantic prosody, and the prospective event may be subverted into a device. In short, a transition in a text will contain a prosody, normal or fractured, that has a special influence on the text's meaning.

To recapitulate, this combination of words may be a lexical collocation, and that is when we use the term ***semantic prosody*** to describe the tendencies of certain lexical items to co-occur and thus create a certain overall tendency of meaning. It may also be a grammatical string, with a certain semantic aura in the reference corpus of which the reader will be unaware, but which will influence the text's meaning. Semantic prosodies of grammatical strings are also opaque to intuition: Louw (2010b) calls them ***logical semantic prosodies***. Or, we may look at the lexis contained in a grammatical string. This lexis may either fit in with the string's normal prosody, or it may deviate from the lexis that is normally used in its slot or slots. The most frequent lexis used in a specific slot within a grammatical string is called its ***subtext*** (see Section 2.6.2). The subtext will still be there, as

a sort of aura, even if the lexical item selected by the author is unusual. Then the event will develop unpredictably and may become a device, depending, among other things, upon the degree of collocational and contextual mismatch. The following sections will argue that collocational mismatch alone is insufficient for the transition to be pronounced a device, and provide evidence for this in English and Russian. This conclusion is in accordance with Louw's (1993) notion of breaches of semantic prosody developing either as irony (we may add, any rhetorical device), or insincerity (ranging from only partial willingness to reveal one's genuine attitude to the issue discussed, to a desire, conscious or unconscious, not to reveal one's genuine attitude).

5.1.4 Relexicalisation and literary devices

Louw has claimed that 'all literary devices have a corpus-accessible feature in common: relexicalisation' (Louw 2008). The definition of ***devices*** accepted here will be that of figures given by Abrams and Harpham (2009: 118):

> Figurative language is a conspicuous departure from what competent users of a language apprehend as the standard meaning of words, or else the standard order of words, in order to achieve some special meaning or effect. Figures are sometimes described as primarily poetic, but they are integral to the functioning of language and indispensable to all modes of discourse.

Relexicalisation is the mechanism through which a word which is used delexically (figuratively) regains, for a moment, its full lexical meaning in the presence of another collocate (see Section 2.5.1). An example could be an article published by *The Times* on 11 December 1995 about the forthcoming BBC Christmas radio show:

> Half a million BBC Radio 1 listeners who tune in on Christmas Eve will hear a new version of the Nativity. The Gospel according to the pop station will feature 'Joey' for Joseph and portrays Mary as a moody character [...]

> The play, in five three-minute episodes, has been written in the strip-cartoon style of Judge Dredd and nicknamed 'Judge Jesus' by those working on it.
>
> Mary blurts out 'Oh my God!' when confronted by the Angel Gabriel […]
>
> He [Joseph] says 'I'll be a laughing stock. I am a descendant of King David, for God's sake.'
>
> She says he is a 'crap' carpenter who cannot make a decent birdbath. 'Mary mother of God!' she says, adding: 'Oh no that's me.'

Expressions containing the word 'God', such as 'Oh my God', have become delexical – they do not invoke the image of the Christian God. The foundation of humour in this radio show is the fact that these expressions are relexicalised – as they are uttered by the participants in the Nativity. The quoted passage is useful as it shows relexicalisation being 'abstracted at the level of syntax' (as in Firth's view of collocation), and not only restricted to Sinclair's 9-word window.

If, according to Contextual Prosodic Theory, all devices relexicalise: 'Where there is no relexicalisation, there can be no device' (Louw and Milojkovic 2016: 136), what do we then make of fractured prosodies that do not contain relexicalised lexis? Do we still regard them as devices, that is, figures of speech, or are they simply telling, give-away signs of an authorial attitude? Finally, if the fractured prosody cannot be detected without resort to corpora, how consciously or intentionally was this attitude expressed?

It is the privilege of the corpus stylistician to start from raw data and then arrive at generalisations, if such are scientifically possible. Nothing at all needs to be or may be stated outside a chunk of text studied in its context of situation and the access to a reference corpus with its patterns in context. Fracture is a matter of degree. Corpora help to display this on a cline. However, common sense suggests that if a fractured prosody cannot be detected without corpus means, the chances are that the author was unaware that his or her attitude had found its way into the text. The practical examples in the next sections may clarify the issues outlined above.

5.2 A prosodic clash in Pushkin

Having shown that semantic prosodies subsist in Russian (and Serbian) in a similar way to those in English, I now propose to illustrate the implementation of prosodic clashes in Russian by means of corpus stylistics. To this end, let us take Pushkin's famous poem 'Друзьям' ('To Friends'). The second stanza contains the expression *оживил войной* - 'enlivened through war'*:*

Его я просто полюбил:	I simply grew to love this man:
Он бодро, честно правит нами;	He rules us cheerfully and honestly;
Россию вдруг он ***оживил***	He enlivened Russia all at once
Войной*, надеждами, трудами*.	Through war, through hopes, through labours.[18]

Although the Russian verb *оживить* may be translated as 'revive', the English 'enliven' reflects the sameness of the root in the Russian verb and the noun 'life' (*жизнь*). Also, in my translation 'enlivened' is separated from 'war' by 4 words, whereas in the original the verb is immediately followed by the noun in the instrumental case, without a preposition. For this reason, I will refer to the collocation as 'enlivened through war', to emphasise the impression of illogicality it creates in Russian.

The reason this famous quote in the poem praising the Tsar is included is not the clash itself, obvious 'with a naked eye' and not requiring corpus data – although, of course, empirical data are always welcome. Nor could we dare to choose a less obvious prosodic clash before thoroughly investigating whether it applied as long ago as the year 1828. The more intriguing question is whether the clash is caused by (conscious) irony or (subconscious) insincerity (Louw 1993). And here it will help to consult the poet's own corpus.

5.2.1 The analysis of *оживил войной*

When the verb *оживить* ('enliven', 'revive') was co-selected with a noun in the instrumental case in the RNC, the corpus (accessed in November 2013) yielded 308 entries.

[18] Translations offered in this dissertation are mine, unless stated otherwise.

Among them there were 5 expressly negative collocates of the verb: *бредом* (‘drivel’), *непогодой* (‘storm’), *слезой* (‘a tear’), *убийством* (‘murder’), and *уколами жадности* (‘pricks of greed’), the last two found in texts by M. Gorky. The most intriguing one at this stage belonged to none other than Pushkin himself: *‘Места, еще недавно оживленные присутствием 15.000 войска, были молчаливы и печальны.’(*‘The area, until recently enlivened by the presence of an army of 15 000, was silent and gloomy.’) When *оживить* was co-selected with *вой** (the root of the Russian noun ‘war’), distance 1-5 words, in the whole of the RNC there were 2 sentences, one where the verb and the noun were not syntactically related, and the other being a quote from the stanza under discussion. When the verb was co-selected with the related root *вое** (introducing an adjective related to ‘war’, or a noun meaning ‘soldier’), there were no matches. In the whole of RNC, lexical items related to the word ‘war’ are never co-selected with *оживить* (‘enliven’, ‘revive’).

When the RNC subcorpus of Pushkin’s prose was consulted, it became clear that at least where his prose is concerned, war (the selected node was ‘война’) is viewed by him not so much as a bloody enterprise but as a patriotic and strategic one, as well as a means to advance one’s career. The sacrifices accompanying it, such as death and destruction, are not presented in detail. (For example, a characteristic line is: *‘В молодых еще летах он успел уже отличиться на поприще войны и гражданственности’ (*‘When still in his youth, he had already distinguished himself in the arena of war and citizenship’, from *The History of Pugachev*).

However, Pushkin’s poetic subcorpus revealed a different picture. This is the second page of the concordance lines yielded by the RNC, here presented in full:

1 Он **удручен** годами, Войной, **заботами, трудами;** Но чувства в
2 Друзья **кровавой** старины Народной чаяли войны, Роптали, требуя **кичливо,** Чтоб гетман
3 Теперь бы **грянуть** нам войною На **ненавистную** Москву!
4 жертве, ласки расточали; И **озабоченный** войной, Гнушаясь мнимой клеветой, Донос оставя
5 плоды страстей, войны, **трудов, Болезни, дряхлость и печали**
6 Незапно Карл поворотил И перенес войну в Украйну.
7 **Пламя пышет,** Встает **кровавая** заря Войны народной.
8 земного В **трудах державства и войны** его товарищи,сыны: И Шереметев
9 нами; Россию вдруг он **оживил** Войной, **надеждами, трудами.**
10 Наслажденья, Как принимал во дни войны Он вызов **ярого** сраженья.
11 над **честью** и над ним, Войну ведет **неблагодарный** С своим **союзником**
12 **любила** [Все роды **славы:**] дым войны И дым Парнасского кадила.
13 На Русь ли вновь идет войною, Несет ли Польше свой закон
14 **бранной славой; Устала грозная** рука; Война от мыслей далека.

15 **Обезображенный** войною, Цветущий край **осиротел; Исчезли** мирные
16 Гирей для мирной неги Войну **кровавую презрел, Пресек ужасные набеги**
17 **Опустошив огнем** войны Кавказу близкие страны И **селы**
18 **печали** о **сыне, падшем** на войне.
19 Другой пускай поет [**героев**] и войну, Я скромно **возлюбил живую тишину**
20 носился, во дни былые, дни войны, На западе, на юге бился
21 Война. Война! Подъяты **наконец**, Шумят знамены
22 Война. Война! Подъяты **наконец**, Шумят знамены
23 Рожденные в снегах для **ужасов** войны, Там хладной Скифии **свирепые** сыны

Figure 5.1 A section of the concordance of *война* from Pushkin's poetic subcorpus of the RNC

The collocates in bold help to establish Pushkin's attitude to war. What follows is the same concordance translated into English (not a concordance of translated texts that make up the Russian version). In order to show the reader unfamiliar with Russian what a Russian reader sees when perusing the concordance, I have translated it in prose, keeping in mind the wider contexts of the poems from which the lines came:

1 He is **depressed** by age, War, **cares**, **labours**; but feelings in
2 the **blood-stained** past wished for the people's war, and grumbled , **haughtily** demanding
3 The time is ripe to **wage a sudden** war on **hateful** Moscow!
4 victim, showering caresses; and **worried** by the war, loathing the sham slander, having
5 the fruits of passions, war, **labours**, **Illness**, **decrepitude and grie**
6 Charles turned his troops and transferred the war to Ukraine
7 **Flames** are **blazing**, The **bloody** dawn of people's war is rising.
8 earthly lot In the **labours of statesmanship** and war his comrades, sons: Sheremetyev
9 us; he **revived** Russia all at once through war, through **hopes**, through **labours**.
10 pleasures, as he accepted in the days of war the challenge of a **violent** battle.
11 onour and himself; **ungratefully** he is waging a war on his **ally**
12 **loved** [all kinds of] **glory**: the smoke of war and the smoke of Parnassus' censer.
13 Whether he be waging another war on Russia, or extending his rule to Po
14 **military glory**; the **formidable** arm is **tired**; war is far from his thoughts.
15 **Disfigured** by war, the blossoming region is now **orphaned**
16 e sake of peaceful pleasures, **abandoned bloody** war and **put an end to his plundering r**
17 **Ravaged** by the **flame** of war the regions and **villages** close to
18 **grief** for her **son**, **who fell** in a war.
19 Let another sing [the **heroes**] and the war, I modestly have chosen **living silence**
20 galloped, in the old days, the days of war, fought in the West and in the South
21 To war. To war! Lifted **at last**, the banners s
22 To war. To war! Lifted **at last**, the bannesr swish
23 Born in the snow, for the **horrors** of war, the **savage** sons of Scythia

Figure 5.2 The English translation of Figure 5.1

The collocates given in bold yield a fairly good account of the poet's attitude to war.[19] It is a bloody pursuit, causing worry to statesmen and desolation to contested territories, but an honourable one. Since collocates thus obtained should by no means be taken at their face value, wider contexts for these lines were opened to discover, for example, the following lines: '*Война! Подъяты наконец, Шумят знамены бранной чести! Увижу кровь, увижу праздник мести; Засвищет вкруг меня губительный свинец. И сколько сильных впечатлений Для жаждущей души моей!*'('To war! Lifted at last, the banners of martial honour swish! I will see blood, I will see the celebration of revenge; the deadly lead will come whining around me. And how many powerful impressions for my craving soul!')

This summarises Pushkin's attitude to war. The expression *оживил войной* ('enlivened through war') in the poem '*Друзьям*' is uniquely characteristic of the poet, but consistent with his view of the world as expressed in the literary world of his poems. There is no (conscious) irony. There is no (subconscious) insincerity either. A loyal subject to the Russian Empire, addressing the Russian Emperor, he maintained that honour was meant to be defended, if need be, even through the deadliness of bullets. If there is a conflict in such a stance, he lived to support it to the last.[20]

5.2.2 Authorial intention in *оживил войной*

What we have discovered in the literary world of this particular poem is overprovided context of situation, (see Louw 2000). If in the whole of the reference corpus

[19] It is worth noting here that the third and last page of the same subcorpus, composed mainly of lines from his earlier work, gives collocates pointing at a more enthusiastic attitude to war dangers. The Poetry Corpus of the RNC has been expanded since then; the impact of the Pushkin concordance was no different in November 2013.

[20] The empirical findings are supported by Yury Tynyanov's book *On The Journey to Arzrum:* 'If Pushkin's attitude to the 1828-1829 campaigns is expressed by the formula:

He enlivened Russia all at once

Though war, through hopes, through labours

then the view of the 1828-1829 imperial wars as of an affair 'governmental and not patriotic' is characteristic of Pushkin also, in the sense that he disagreed not so much with the war as with those who were waging it and with how it was waged.' (http://az.lib.ru/t/tynjanow_j_n/text_0110.shtml)

only Pushkin's text co-selects 'enliven' and 'war', then the discrepancy between the reference corpus and the text points to overprovision, in the same way in which Philip Larkin's references to love, mainly defeating hope, point to the underprovision of context (Milojkovic 2011a).

I believe that I have uncovered Pushkin's authorial intention: that he meant what he wrote, without reservation. In fact, here the first collocate hinges on the second collocate in its meaning of 'campaign', rather than 'war', as stated by Yury Tynyanov. The collocates 'war', 'hopes', 'labours', in the eyes of the author, share a constructive, proactive semantic prosody. ***Authorial intention*** is thus here defined as the intention of an author to convey a certain meaning.

5.2.3 Relexicalisation, authorial attitude, and device

So far we have encountered one prosodic clash in Russian – 'enlivened through war'. Earlier in the text it was hypothesised that if all devices relexicalise, it must mean that, in the absence of relexicalisation, a prosodic clash is a tell-tale sign of authorial attitude rather than a device. What I had taken for a tell-tale sign of scepticism in Section 2 turned out to reveal Pushkin's 'logical construction of the world' (see Section 2.6.1). Apparently, in the days of Pushkin, wars were viewed as military campaigns, and not as a national tragedy, which is what they are considered today. In this case, the Russian verb *оживить* ('enliven', 'revive') is used delexically in the text, in the meaning of 'energise', 'inspire'. This is supported by context clues. For example, the following 4 contexts in the main corpus of the RNC confirm that hope enlivens (I searched for *надежда* ('hope') preceeding *оживить* ('enliven'), 1 to 5 words intervening):

1 ***Надежда оживила*** *его, он вцепился в Кривицкого, упрашивая его сейчас же сходить на склад. [Даниил Гранин. Искатели (1954)]*
(Hope enlivened him, and he clung to Krivitsky, begging him to go to the warehouse at once. Daniil Granin,The Seekers)

2 Его энергия, его бодрость, его ***надежда оживили*** *и всех остальных. [П. И. Ковалевский. Петр Великий и его гений (1900-1910)]*
(His energy, his cheerfulness, his hope enlivened everyone else. [P. I.

Kovalevsky. Peter the Great and his Genius 1900-1910)]

3 В ее груди вдруг болезненно ярко вспыхнула все время незаметно тлевшая ***надежда*** *и* ***оживила*** *ее… [Максим Горький. Мать (1906)]*
(In her breast suddenly, painfully and sharply, broke out a flame of hope, which had been hitherto smouldering there, and that flame revived her… (Maxim Gorky. Mother. (1906)])

4 *Он произнес эти слова с такою уверенностью, что* ***надежда*** *снова* ***оживила*** *молодую женщину. [К. М. Станюкович. Нянька (1895)]*
(He spoke these words with such certainty that hope again revived the young woman. [K. M. Stanyukovich. Nanny (1895)]

Figure 5.3 The RNC contexts of *надежда* ('hope') preceeding *оживить* ('enliven'), 1 to 5 words intervening (with provided English translations)

Hope enlivens, but it never literally raises from the dead, as shown by the states of affairs in the 4 contexts. The generally delexical use of 'enliven' was confirmed by the RNC. Russia as 'enlivened through war' is a prosodic clash if the verb is understood literally, because the verb shares its root with the Russian noun signifying 'life'. Still, all the other collocates point to the verb's delexical use. This means that relexicalisation occurs within the collocational mismatch, but not outside it. This fact has a bearing on the existence of devices in texts.

The degree to which a fracture is a device may depend on the degree of the relexicalisation of its lexis. In the present case, the degree of relexicalisation (etymolohy) is overridden by the existing delexical use and the positive context clues. Still, the question remains: what makes a movement on the cline between the delexical and the relexicalised worthy of being called a device? This question can only be answered fully with the involvement of subtext, as will be illustrated in the subsequent sections.

5.3 Delexical subtext vs. subtext relexicalised: which one is part of a device?

We have seen that relexicalisation is prompted by proximity alone in the form of collocation. The weakest form of collocation is purely lexical. Its strongest form collocates

grammar strings with necessary forms of metaphysics (lexis) and its functions. The reason why the latter form is the strongest is its impenetrability - as far as mere intuition is concerned (Louw and Milojkovic 2016: 189-190). If I spotted the illogicality of 'enlivened through war' without recourse to corpora, I cannot hope to do the same when it comes to grammatical strings collocating with lexis embedded in or around them.

In short, so far we have considered relexicalisation, which is realised, as its very name suggests, through lexical collocation. Now we are about to examine another form of collocation, in which grammatical strings collocate with lexical items embedded in them. As explained in Section 2.6.2 and reiterated in Section 5.1.3, the subtext of any grammatical string consists of the most frequent lexical choices made in the slot occupied by the lexical item. In CPT they are called quasi-propositional variables (QPVs). Of course, only a large reference corpus will show us what lexical choices are made within a certain grammatical string, and which are more frequent than others.

The remaining part of the section will propose a mechanism for establishing whether a certain lexico-grammatical combination (a) is a device, and (b) conveys the author's intention. Both these questions will be answered by applying the notions of delexicalisation and relexicalisation to the lexical variables within and around grammatical strings. In order to illustrate this mechanism, I will study the subtext in two different English examples, and then apply the findings to an instance of subtext in Pushkin.

In the poem 'Daughter of the House' by Herbert Williams, the unmarried female occupant of a house feigns dedication while hiding her opportunism (Louw 2010a: 93, see also Section 2.6.2). This is revealed by the subtext of the line 'a heart made bleak by sacrifice'. The searchline 'a * made * by' reveals, in the second slot, the adjective 'possible' in the place of 'bleak' as the most frequent lexical choice. The heroine is, below the surface of self-sacrifice, opportunistic. This is the subtext of the line as well as of the whole poem. Herbert Williams himself confirmed, in correspondence, that this was indeed the impression he intended to convey (Louw 2010b). Williams also expressed surprise at his intention being discovered. This is the poem:

It is not love that keeps her here, tending
The stubborn enterprise of age. Her hands
Are clinical expressions of *a heart*
Made bleak by sacrifice, her eyes
Neutralise her therapeutic smile.

Love is an easy master, but her **guile**
Springs from more terrible demands.
It is the blood's dictatorship, bending
Her uninvited kinship to the **part**,
Masking indifference with a **knack of lies**.

In the second stanza , the words 'guile', 'part', 'masking' and 'knack of lies' support the most frequent QPV 'possible' (as in 'make possible') and Louw's interpretation. The poem shatters the usual images of the young as either sacrificing themselves out of filial duty or being suffocated by it. The woman who lies and plays the part may well turn out to be opportunistic, even if led to it by a need for suppression. Either way, we could not have known what the subtext of the string is without recourse to corpora, whether it is at odds with the persona's reasoning, and whether it is supported by the context clues. Subtext is opaque to intuition.

It is possible to say that 'guile' and other context clues pointing to opportunism bring to life the hidden subtextual meaning. It is, then, an instance of relexicalisation of the subtextual lexical variable. This leads to the conclusion that if the QPVs of a grammatical string are relexicalised or otherwise supported by the context clues, the meaning is expressed as intended. This answers question (b) posed above. As to question (a), in this case 'bleak' was not intended to serve as a literary device conveying the intended meaning because the support given to 'bleak' by the relexicalisers stays at the level of the delexical. This will be shown below.

A 'bleak heart' is delexical. The BNC does not contain a collocation made up of 'bleak' and 'heart'. COCA contains one example, describing a new widow: 'Black clothes,

black humor, bleak heart.' Despite this paucity of examples, the collocation is not infrequent: the Google Books corpus offers 80 contexts where 'bleak heart' is used delexically, in the meaning under discussion and in similar contexts of situation. This is a concordance of the first 20 contexts yielded by the corpus:

1 hort time, used as a hyphenated name Bleak-Heart, for business purposes, it (2010)
2 he anguish and remorse that fills my bleak heart it alone can know. Yet, could 1842)
3 rs, by he cold world entombed in his bleak heart, had petrified. And is Louise 1849)
4 , at 'old girls', but this betrays a bleak heart and bad taste. Matured maidens (1859)
5 al Progressive party, he went with a bleak heart. 'It is a burden greater than any (2013)
6 r, that now felt as empty as his own bleak heart. Zoe wondered up the marble steps,(2010)
7 s peace. All was peace except Lucy's bleak heart. Bluey woke her. The pup was let in(2011)
8 or letting go, forgiveness blue, the bleak heart's ancient ache for summer things,(1987)
9 Taut posture Cloudy curves Hazy head Bleak heart Stoic stomach Lean legs. (1995)
10 king scene, she experienced only a 'bleak heart'. Walking down a long narrow (2004)
11 tterfly. But in the sad lore of his bleak heart, the father read the meaning of (1918)
12'It's not fair – it's not fair!' her bleak heart was crying out while she tried to (1916)
13 o quiet the troubled mind, O'er the bleak heart breathed with a spirit bland, Like (1888)
14 om grief hath driven mad – in whose bleak heart All natural affections have died (1876)
15 some sunlight in time into her own bleak heart. But she said nothing to Isom, and (2012)
16 s had robbed her of the thing her bleak heart craved more than it could ever crave (1910)
17 ility. She yielded to what, in her bleak heart of hearts, she had to do. She bowed (2014)
18s flock', from the Messiah. Many a bleak heart has been thawed by the simple ballad (1916)
19 dance? What grudging spirit, what bleak heart conceived those mockeries of resting (1900)
20 dered to myself if a desolate and bleak heart can ever find a place to start living(2011)

Figure 5.4 The first part of the concordance of 'bleak heart' from the Google Books corpus

In every context in the concordance 'heart' is used delexically. 'Bleak' in this combination cannot be understood otherwise than figuratively. Therefore, as a collocation, 'bleak heart' is made up of two delexical collocates. However, it is also worth noting that 'heart' is relexicalised in its meaning of a human organ through the collocates 'hands', 'eyes', smile', which are all physical. As for 'bleak', it is not relexicalised. In order to become so, 'bleak' has to regain the quality of being cold and gloomy, as a landscape or weather conditions (these collocates appear in the BNC when 'bleak' is used in its full lexical meaning). There are no relexicalisers that bring this meaning to mind. 'Neutralise', in the next line, does in part support the meaning of 'bleak', but sounds abstract enough.

Weather is never neutral or neutralised (this is confirmed by the BNC and COCA); landscapes might be described as such (no such examples in the BNC or Google Books; there are 2 instances in COCA only, and in one 'neutral' appears in inverted commas, implying the rarity of the collocation). 'Neutralise' is therefore a poor relexicaliser. 'Clinical' supports the use of 'bleak' without fully relexicalising it either. Thus, the epithet 'bleak', which is generally used delexically more often than literally, and collocating with the delexical meaning of 'heart', is not properly relexicalised – 'except by its collocate' heart', which is relexicalised through 'hands', 'eyes' and 'smile', used literally. We may conclude that the sequence 'a heart', used delexically but relexicalised, is an intended device. The proof that the device is intended is the visual grouping of 'hands', 'heart', 'eyes' and 'smile' to the right, each ending a verse. The parallelism is intentional. On the other hand, the sequence 'made bleak by sacrifice', in which the subtext is supported by context clues, but the lexical item poorly relexicalised, was not meant to be a literary device by the poet, although it does express the poet's intended meaning, or authorial intention. This answers question (a) posed at the beginning of this section. To conclude, 'made bleak by' describes opportunism, inadvertently in this particular string, but this generally being the author's intention. This was not intended as a literary device, unless viewed in a wider co-text in collocation with 'heart', the use of which was intended as a device.

It stands to reason that any authorial meaning expressed through a grammatical string is expressed unintentionally, since the subtext of a grammatical string is opaque to intuition. However, much depends on the lexis inside or around a string. Let us consider our other English. In another poem by Yeats, 'The Circus Animals Desertion', in the coda, the persona says: 'Now that my ladder's gone,/ I must lie down where all the ladders start'. The most frequent lexical choices in the place of 'ladder' in the searchline 'now that my * is gone' are either very significant persons, or qualities without which a decent life is an impossibility (Louw and Milojkovic 2014: 279). However, the relexicalising collocates (relexicalisers) which immediately precede the quoted lines are those making up the literary world of a rag-and-bone shop. They bring to life the palpable quality of a ladder as an object one might easily find in such a place. These relexicalisers are what makes the

‘ladder’ a metaphor – a device. The ladder sounds as if it were physically present in the shop but at the same time represents the persona’s most valued person and/or greatest personal value. Their being gone is the cause of utmost despair that only the subtext of the line can show.

The relexicalising power of the collocates restores the physical aspect of the ‘ladder’, while the subtext brings to life its abstract and even ‘human’ qualities. The whole cline, from the very literal to the highly relexicalised, including the middle ground of the traditional view of the ladder as a parable of spiritual advancement, has been covered at one stroke.

What is the difference between ‘a heart made bleak by sacrifice’ and ‘now that my ladder’s gone’? It has to do with (1) the degree to which the lexical collocate of the grammatical string is delexical in terms of lexicography and (2) the degree of its relexicalisation by the surrounding collocates.

The difference between the subtext in Williams’s poem and in Yeats’s ‘Circus Animals’ lies in the nature of relexicalisation. While context clues in the second stanza in Williams support subtext and stay at the delexical end of the delexicalisation-relexicalisation continuum, relexicalisation in Yeats supports the physical aspect of the lexical collocate of the grammatical string while the subtext of the string speaks of its lexical collocate’s utmost spiritual relevance to the persona. Neither poet could have been aware of the subtextual meaning of the grammatical strings they employed. However, Yeats managed to broaden the scope of the metaphor’s associations from base certainty to elevated despair, while Williams stayed at the spiritual, delexical end of the scope. In other words, Williams’s subtext is further delexicalised by the surrounding context clues, and Yeats’s is truly relexicalised. Hence the claim that ‘now that my ladder’s gone’ was intended as a literary device, while ‘a heart made bleak by sacrifice’ is a fractured prosody which does convey the author’s intended meaning, but in which the use of ‘bleak’ was not intended as a device in the way that the ‘ladder’ was – because it is poorly relexicalised. This is in accordance with Abrams's definition of figurative language as 'conspicuous departure from the standard meaning of words' (see Section 5.1.4). However, ‘made bleak by’ is an example which also goes to show that on the ‘literal-delexical-relexicalised’ cline

of meaning, there are movements in between the delexical and the relexicalised that make devices difficult to define, and that exclude sweeping generalisations. The 9-word window is dynamic, and meanings interact both within the text and intertextually.

In this section, subtext was illustrated through two different poetic examples. These examples present a cline as to the degree to which (a) the lexical choice of the poet in the given grammatical string is relexicalised in the co-text, and (b) to which the subtext of the grammatical string is relexicalised or otherwise supported by context clues. As previously stated, the CPT hypothesis suggests that (a) all devices relexicalise, and that (b) authorial intention in a poetic text will be supported by the surrounding context clues, because poetic texts are well thought-through. In order to compare the two presented strings containing the lexis of the poets' choice, the findings are summarised in the table below. The strings in question are viewed against the background of Yeats's 'that is no country for old men', discussed in the Louw (2010b, see also Section 2.6.2.3):

Table 5.3 Literary devices and authorial intention established by means of relexicalisation and context clues.

text/ lexical choice, subtext and context clues	1 Is the author's lexical choice relexicalised?	2 Do the context clues support the subtext of the grammatical string?	3 Is the lexico-grammatical collocation a device?	4 Does the lexico-grammatical collocation express authorial intention?
'that is no *country* for'	no, it is used in its full lexical meaning	no, the context clues do not support 'reason' or 'excuse'	no	no
'(a heart) made *bleak* by'	poorly	yes	no	yes
'now that my *ladder*'s gone'	yes	yes	yes	yes

A literary device, then, will be contained in a grammatical string whose subtext is supported by context clues, co-selected with relexicalising lexical items. Therefore, for a lexico-grammatical collocation to be pronounced a device, questions 1 and 2 in Table 5.3 will be answered affirmatively. For establishing authorial intention, it is enough to find that the author's subtext is supported by context clues – therefore, it is enough to answer question 2 in Table 5.3 affirmatively. In Section 5.4, a grammatical string from a stanza by

Pushkin will be analysed in order to establish whether it contains a literary device and/or expresses authorial intention.

5.4 Establishing subtext in Russian

What follows is a detailed subtextual analysis of a grammar string from the first stanza of one of the most celebrated love poems by Alexander Pushkin. It will attempt to disclose the additional layers of meaning in the stanza, using the principles explained in the preceding sections of this section. In line with the present argument, the studied line will be included in Table 5.3, after the questions of the presence or absence of a literary device in the grammar string and of authorial intention in the string's subtext are answered.

5.4.1 The author's collocation and the evidence

If there are clues in Williams' poem which support the subtext of 'made * by' – we found them in the second stanza – that is not the case with Pushkin's stanza I am about to quote. The example in question is the poem by Pushkin (Milojkovic 2011b) in which the subtext of the second line of the first (and also the penultimate) stanza goes very much *against* the apparent meaning of the stanza and, indeed, the whole poem. Yet, the subtext is there. Do we simply say that it is atypical use, or do we state that it was not the author's intention to convey the meaning expressed by the verse's subtext?

This is the beginning of one of the most celebrated love poems in the Russian language:

Я помню чудное мгновенье:	I remember a wondrous moment:
Передо мной явилась ты,	**You appeared before me**
Как мимолетное виденье,	Like a fleeting vision,
Как гений чистой красоты.	Like a genius of true beauty.

Here is the English translation by J. E. Falen (Pushkin 2009: 89):

I still recall a wondrous vision:

That day when I beheld your face,

A fleeting moment's apparition
Of perfect beauty and of grace.

Whenever hapless grief oppressed me
Amid life's cares and pointless schemes,
Your gentle voice and soul caressed me,
Your cherished features filled my dreams.

The years went by. Fate's storms and stresses
Dispersed those sacred dreams of grace,
And I forgot those soft caresses,
Your gentle voice and angel's face.

In bleak despair and isolation,
My days stretched on in quiet strife:
No awe of God, no inspiration,
No love, no tear, no sense of life.

And now once more I've seen that vision:
My soul awoke; I saw your face,
A fleeting moment's apparition
Of perfect beauty and of grace.

My spirit soars in exultation,
And once again there reappears
The awe of God… and inspiration…
The sense of life… and love… and tears.

Let us return to the original. Pushkin's chosen word order in the second line is poetic and the exact grammatical pattern yielded no results in the RNC. When I reversed the word

order, the search line *ты *сь перед** (‘you (thou) *ed before’) gave the following 6 lines (the first one being extremely trivial):

А чем ты умывалась перед нанесением крема?
„Давай поиграем“, - предлагаешь ты, **прохаживаясь** перед ним в соблазнительном
помню один чудесный момент, когда ты появилась передо мною, как будто
Помнишь, ты **извинилась** перед гостем, вышла за
- Чем же ты **провинилась** перед волшебницей, что она
так испугала моего садовника, зачем ты **появилась** перед ним?

And what did you wash your face with before applying the cream?
‘Let’s play’, you suggest, strutting before him in seductive
Remember a wonderful moment when you appeared before me as if
Do you remember, you apologized to a guest, came out of
How have you offended the sorceress so she should
righten my gardener so, why did you appear before him?

Figure 5.5 The RNC concordance of *ты *сь перед** (‘you (thou) *ed before’), with its English line-by-line translation

The grammar of Russian requires that certain verbs, e.g. *извиниться* (‘apologise’), *появиться* (‘appear’), *провиниться* (‘wrong’, ‘offend’) be followed by the preposition *перед* (‘before’) if the indirect object of the verb is mentioned. At first, this is how I interpreted the concordance above – that the reference corpus simply yielded several verbs whose grammatical behaviour required this preposition. Then I noticed that the contexts of situation pointed at either guilt or sexual provocation. There are 2 exceptions. Line 3 is a paraphrase of Pushkin’s famous line. Line 1, its triviality aside, employs ‘before’ in its temporal sense: in Firth’s terms, a relevant person/personality is missing. In 2011 I drew the following conclusion: ‘The third line is an intentional paraphrase of Pushkin. The last 3 lines create contexts of guilt. The second line creates a context of sexual provocation. The sample is almost too small for a tentative attempt to interrogate the poet’s attitude to the addressee, but the quasi-propositional variables of guilt and provocation may be pointing at suppressed mixed feelings of the poet. A larger reference corpus is necessary to provide better access to subtext’ (Milojkovic 2011b).

One might state that the RNC is not big enough a corpus and leave it at the very tentative conclusion that had more lines been found and had they more or less confirmed to this pattern, we might be justified in suggesting that the QPVs point to the 'relevant personality's' – the woman's – guilt and provocation. However, Pushkin's biographical data fit into this interpretation so perfectly, and the tone of the poem is so unlike the subtext of the chosen string, that one is tempted to investigate further.

Apparently, Anna Kern, to whom the poem was dedicated, was an extremely attractive young woman, married to an old military general against her will and at a very young age. The marriage was a failure, as the wife's temperament was not suited to obeying her marriage vows in such circumstances. We know from Pushkin's letters that his affair with her took place a year and a half after the poem was written; the affair itself, if it was long enough to be termed that, he describes in a very basic language. In another letter he calls her, jocularly, 'our whore of Babylon'. All this stands in sharp contrast with the poem's adoring tone, but if we imagine a married woman, young and very attractive, appearing 'before' Pushkin, we are tempted to marvel at the precision of corpus findings which have given us a peep into Pushkin's state of mind.

There is no reason for biographical data to be completely ignored; they are the macro-context of situation. But we can either prove the existence of subtext scientifically or leave it for better days. Therefore I will focus on the second line of the first stanza, using the corpus means available, and only those. Are we dealing with atypical use or authorial insincerity? Is there a hidden meaning to the famous poem?

Atypical uses exist because there are typical ones, and both are contextual instantiations of language patterns. In the case of Yeats, I claimed that the string 'Now that my ladder's gone' was a device, because of the presence of relexicalisation which brought to life the ladder as a palpable object. I also claimed that the string expressed the meaning intended by the author, because the subtext was supported by the context clues. In the case of Williams, I claimed that 'a heart made bleak by sacrifice' expressed the meaning intended by the author but was not a device. I justified this by saying that there are clues in the text that support subtext, but there are no relexicalisers of the literal meaning of 'bleak' – rather, several collocates relexicalise the literal meaning of 'heart'. In Pushkin's case

there is a relexicaliser, ‘fleeting vision’, which supports the literal meaning of ‘appeared’, and it will be shown in the following section. Still, nowhere in the whole poem do we find support for the subtext of either female guilt or female provocation – unlike in Yeats’s ‘Circus Animals’ Desertion’, where the clues in the whole poem support the subtext of ‘now that my ladder’s gone’. Moreover, the whole poem is suffused with adoration. In the case of Pushkin, therefore, we find relexicalisation embedded in subtext which contradicts it, and flouts intended meaning. There is no conscious authorial intention that the apparent meaning be flouted.

What were we to make of Pushkin’s subtext if it were to prove conclusive? Would we say that the true meaning of his verse(s) was in fact contempt and not admiration? This would be a huge over-simplification, not taking into account the whole of the poem in which the verse appears. But authorial meaning is everything the text is made of, the maximum of information on any unit of language yielded by reference corpora, enhanced by the information of how this maximum is embedded in a text by a particular author. Subtext has as much right to be studied as part of the text as any other of its parts.

Before we proceed to discuss the subtext in Pushkin in more detail, we need to tackle two questions. The first is, since subtext is made up of a string’s most frequent lexical collocates, does every string always mean the same (its most frequent computer-recovered lexical collocates always being the same)? The second question is, how much control does the author exercise over subtext?

The answer to the first question is – yes. In theory, all grammatical strings have the same subtext. However, this does not suggest ‘uniformity of meaning’ (Hunston 2007: 250), as the relationship between the QPVs and the author’s lexis is always unique in the given context of situation. Moreover, grammatical strings come in shorter and longer varieties, every added grammatical word creating (slightly) different, more refined subtext. Also, much depends on the circumstances of the text’s creation. A sentence pronounced in a hurried conversation differs from one in a poem in that its content may not be as carefully thought out; enough time may not have been spent on thinking about it; the feeling that has originated it may not run that deep or have been felt that long. Also, and most importantly, the analyst must take into account the similarity of events. The subtext will be all the more

telling if the context of situation in the studied language event proves similar to those where the frequent quasi-propositional variables (lexical collocates of grammatical strings) are found. Therefore no judgements must be made until every detail has been studied. A finding that looks sensational might mean that an oversimplification is taking place.

As for the question of how much control the author exercises over subtext, the answer is – none, except in the unlikely event of his or her possessing computer-generated theoretical knowledge of a grammatical string's behaviour. We exercise (some) control over lexis; grammar controls us, so to speak. This means that a writer must trust the grammatical string that seems to him or her the most fitting, and that best expresses his or her authorial meaning. No writer will consider their grammatical strings as separate from their lexis; the lexico-grammatical combination that the writer trusts as the best conveyor of his or her truth – or lie – will write him or her as well as be written by him or her. The choice of grammar, as of lexis, is a result of authorial inspiration, but not immediately traceable as to its impact. It identifies the best writing and allows insights into why poor writing is rejected (see Section 9 on inspiration).

Let us now consider the subtext of *передо мной явилась ты* ('you appeared before me') in more detail.

5.4.2 Relexicalisation, co-selection and states of affairs in establishing subtext

How do we know that 'fleeting' and 'vision' are the relexicalisers of Pushkin's lexical choice ('appeared') in the grammatical string under discussion? I co-selected *явиться* ('appear') and *видение* ('vision') in the RNC, in this order, 1 to 5 words intervening, and found 7 contexts in the Poetry Corpus, 2 of which (4 and 5) came from Pushkin's poem under discussion. In the reversed order of the co-selected collocates, I found 4 texts in the same corpus. Out of these 11 stanzas, written between 1817 and 1903, 3 contain the word 'vision' used figuratively, while in the remaining 8 it is used literally. The delexical vision is that of a beautiful woman in all 3 contexts; the literal vision may be uplifting or frightening. This is illustrated by the 3 contexts below:

1 *Ты раз* ***явилась мне****, как дивное* ***виденье****,*
Среди бесчисленных, бесчувственных людей, —
Но быстры молодость, любовь, и наслажденье,
И слава, и мечты, а ты еще быстрей.
[*А. А. Фет. «Напрасно, дивная, смешавшися с толпою...» (1850*)]
(You once **appeared to me**, like a wonderful **vision**,
Among innumerable, insensitive people –
But youth, love, and enjoyment are brief,
So is fame, so are dreams – and you are briefer still.
[A. A Fet 'In vain, beautiful one, mixing with the crowd…' 1850])

2 ***Мне явятся*** *веселые* ***виденья****,*
Мне явятся *далекие друзья.*
[*А. С. Хомяков. Ударил час, прощайте, други!..:«Ударил час, прощайте, други!..»:[При прощаниях, 3] (1828)]*
(Merry **visions** will **appear to me**,
Friends from distant lands will **appear to me**.
[A. S Homyakov, 'The hour has struck, farewell, friends!', from 'Farewells, 3. 1828])

3 *Сарапис снова* ***ему*** *в ночном* ***явился виденьи****,*
Грозно вещая: «Тебе ль благости ждать от богов!
[Д. В. Дашков. Отсроченная казнь : «Ветхую стену опорой избрав, повествуют, убийца...» [Цветы, выбранные из греческой Анфологии, 27] (1827)]
(Again Serapis **appeared to him** in a nightly **vision**,
Prophesying terribly: 'You! to expect favour from the gods!'
[D. V. Dashkov. An execution postponed: 'Upon choosing a crumbling wall to lean on, the murderer…' [Flowers selected from a Greek Anthology, 27] (1827)])

Figure 5.6 Three contexts of *явиться* ('appear') preceding *видение* (vision') in the Poetry Corpus of the RNC

The finding that 8 stanzas out of 11 tell of a vision in the literal sense of the word supports the claim that *видение* ('vision') relexicalises Pushkin's lexical choice *явилась* ('appeared') in the discussed grammatical string.

What interests us here is how the relevant person witnessing the vision gets a

mention. In the contexts where it does, it is mainly in the dative case. The 2 exceptions are Pushkin's stanza under discussion and a poem by Zhukovsky, Pushkin's contemporary. In both there is the preposition *перед* ('before') – in Zhukovsky it is its archaic and poetic version, *предо* – followed by the first person pronoun (referring to a man) in the instrumental case. If Zhukovsky uses the preposition similarly to Pushkin, then our findings might be undermined – two prosodic clashes in two contemporary poems are less likely than a supposition that perhaps what we have taken for a deviation was, for example, less so in the time of Pushkin and Zhukovsky.

The humorous poem by Zhukovsky, at least as well known as Pushkin at the time, is addressed to Countess Shuvalova and describes an episode when she surprises her guests by suddenly appearing from the darkness of the garden wearing the mask of a dead man. The poet jokingly describes his horror at seeing the deathly features instead of the lovely ones of the Countess, scolds her for her thoughtless gesture, but proceeds to comment on the transience of beauty and the uncertainty of hopes. This is the context yielded by the Poetry Corpus:

...предо мной, в одно мгновенье,
На место прелести младой,
Явилось грозное виденье

The poet complains: '... before me, in a moment, /Instead of sweetness young,/ Appeared an awful vision...'

Given that this is the only other context of the 11 in the Poetry Corpus where the states of affairs are chunked by the collocates *явиться* ('appear'), *видение* ('vision'), and *перед* ('before'), it is remarkable that the preposition *перед* ('before') participates in a language event describing a momentary and trivial, but very unpleasant surprise. It is also remarkable that the expression *в одно мгновенье* ('in a moment') resonates with Pushkin's *мимолетное виденье* ('fleeting vision'). Visions never last very long; still Zhukovsky's and Pushkin's contexts of situation are all the more parallel with the presence of this third

collocate, while the meaning remains all the more remarkably opposite:

Передо мной явилась ты,	**You appeared before me,**
*Как **мимолетное виденье**,*	As a **fleeting vision**,
*Как **гений чистой красоты**.*	As a **genius of true beauty**.
...предо мной, в одно мгновенье,	... **before me**, **in a moment**,
На место прелести младой,	Instead of sweetness young,
Явилось грозное виденье	**Appeared an awful vision**

These events are very similar in their contexts of situation: a beautiful woman appearing before a man; however, they are opposite in content. In Zhukovsky the woman is wearing a horrible mask, which, understandably, causes an opposite reaction from that experienced by Pushkin's persona.

The fact that these opposite states of affairs were created by contemporary poets led me to investigate the states of affairs chunked by *перед* (‘before’) and *явиться* (‘appear’), in this order, 1 to 5 words intervening, in the main corpus of the RNC. This step was also justified as I had had to reverse the word order within Pushkin’s grammatical string in order to study it in the RNC.

The main corpus yielded 127 contexts. Arguably, the classification below could have been different coming from a different researcher; but since each group of contexts will be described in detail, replicability of findings related to the relevant aspects of these texts remains assured. The 4 relevant categories of contexts can be found in Table 5.4, along with concise notes and comments:

Table 5.4 Contexts of *перед* (‘before’) preceding *явиться* (‘appear’) in the RNC

1. Woman appearing before man (10 contexts)	**2. Man appearing before woman, in the context of a love relationship (11 contexts)**

a) 1 context of mother appearing to son (Coriolanus); the woman controls the situation **b)** 3 contexts of a woman before a group of men; the woman is inferior in all 3 contexts but at the same time intimidating in 2; the woman does not control the situation **c)** 7 contexts of a woman appearing before a man, who is attracted by her: • Contexts labelled 43 and 98 – the same quote from Tolstoy's *War and Peace*. The appearing Natasha has betrayed Bolkonsky in the past, and has come to apologise to her former and now dying fiancé. • Context 6, from Pushkin's *The Squire's Daughter*: the daughter of a landed gentleman dresses up so as not to be recognised by her lover • Contexts 22, 61, 79 and 96 all present an attractive woman who is viewed negatively (this is shown by but the negative collocates in the wider co-text) ***Comment:*** The (c) category is the most relevant as it is closest to Pushkin's stanza under discussion in terms of the context of situation. All these contexts signify an unexpected transition. Only one context is very positive – the appearance of Natasha in Tolstoy. In all other contexts the woman is viewed negatively and yet she is the one who controls the situation; the man is the passive witness of a birth (Context 61), or change (Context 79), or disguise (Contexts 6, 98).	**a)** 2 contexts where the man appears before his wife, who controls the situation **b)** 5 contexts where the woman controls the situation **c)** 4 contexts where the man controls the situation (including a context from Pushkin's *The Squire's Daughter*, labeled Context 5: a young man assumes the fashionable pose of disappointment in order to fascinate the ladies) ***Comment:*** In the majority of the contexts (7 out of 11) the woman controls the situation. This has a bearing on Pushkin's stanza given that these are all love relationships; however, the grammatical strings are bound to be different due to the masculine ending of the verb.

3. The supernatural (13 contexts)	**4. Miscellaneous (41 context)**
a) contexts related to God and Christ (4 contexts) • 2 contexts where Christ appears, before a saint (Context 1) and a hermit (Context 31), to protect him from an onset of numerous temptations • the appearance of a warrior saint exterminating the enemies of Christ's name (Context 91) • the princess will not appear guilty before God (Context 81, also a turn of phrase) **b)** contexts related to creatures from hell (5); personalities more powerful than persons • 3 contexts where the vision is extremely frightening (Contexts 74, 78 (also sent to punish sins) and 69) • 2 contexts where it is a ghost – Contexts 70, 117. These are underplayed: both refer to *Hamlet* and both are humorous. **c)** Visions and apparitions – 4 contexts (26, 60, 66, 87). Only Context 26 is negative; the rest are positive, even if the vision is frightening (punishing sinners) ***Comment:*** Interestingly, the appearance of the supernatural is most often positively viewed. Even if it is a creature from hell, it is here to punish sinners.	**a)** 19 contexts with negative collocates **b)** 6 contexts with positive collocates (3 contexts of heroism and 2 contexts of the supernatural) • war heroes deserving admiration in 41, 68 • Sakharov, a heroic dissident appearing to the Soviet people in Context 16 • a religious trance (Context 121); a fairy tale (Context 40) in which after magic words anything you wish for appears before you • a film director appearing in a new light after filming 'Idiot' (Context 19) **c)** 16 remaining contexts • 6 ironic descriptions (Contexts 2, 8, 13, 83, 50, 100) • 5 appearances before justice (Context 36; harsh military justice in Contexts 125, 126, 127; God's judgement after capital punishment in 114) • 2 divestments of power (Contexts 77, 94) • a context from Pushkin; sudden military threat encountered on a deceptively peaceful meadow (Context 93) • a sudden appearance in Context 30 • an appearance of masks in Context 102 ***Comment:*** In the miscellaneous contexts there are, in effect, 25 negative states of affairs out of 41 (including the 6 that do not contain negative

	collocates but prove to be ironic). Positive contexts include 1 context of moderate approval and 5 of great admiration or wonder. Overall, if a context is not strictly negative, there will be a relationship formed in which one of the two parties (the appearing and the witness) is likely to be inferior to the other.

As Table 5.4 shows, the states of affairs chunked by the co-selection of *перед* (‘before’) and *явиться* (‘appear’), in this order, yielded 4 relevant groups of contexts that can be discussed in Firthian terms: (1) of a woman appearing before a man, (2) of a man appearing before a woman, (3) of the supernatural and (4) the ‘miscellaneous’ group. Occurrences of *перед* (‘before’) in its temporal meaning and turns of phrase containing *явиться* or *являться* (appear, perfective/imperfective aspect) in formal language were ignored.

How can these findings, describing states of affairs conjured up by 'before' and 'appear' in Russian, be explored in relation to Pushkin's 'genius of true beauty'? The message of Pushkin's stanza and poem is the divine appearance of a woman, compared to a vision, before the adoring man. This corresponds to the following categories in our classification:

1. (c) the appearance of a woman before a man (relevant due to the context of situation)

3. (a) the appearance of divinity of the highest order (relevant due to Pushkin's wording 'genius of true beauty'

3. (c) visions and apparitions (relevant due to Pushkin's wording 'fleeting vision')

4. (b) miscellaneous contexts with positive collocates (relevant due to Pushkin's overall tone)

There is a definite discrepancy between the group of contexts classified as 1 (c) and

the other groups above. In the corpus, women appearing before men are attractive to them, and their attractiveness is a feature which is relevant to the context of situation. However, in each context but one (43, same as 98, the appearance of Natasha Rostova to mortally wounded Bolkonsky) the woman is seen as inadequate. (In Context 6, the heroine enjoys our approval and sympathy but disguises herself so as to seem inadequate). The only very positively viewed appearance is that of Natasha Rostova in Tolstoy's *War and Peace*. However, innocent Natasha has wronged Bolkonsky in the past, albeit acting in good faith, and the wrong she has done him has had a definite effect on the life that he is contemplating at the moment of her appearance. Therefore, in every single context featuring similar participants to Pushkin's context of situation, something is wrong with the woman in the eyes of the man to whom she appears and who finds her, in various degrees but always relevantly, attractive.

On the other hand, the contexts involving divinity, visions and worshipful feelings (3a, 3c, 4b) do not seem in their majority to involve women. In short, Pushkin uses two collocates – 'before' and 'appear' – which, in the reference corpus, create two states of affairs: either of an inadequate woman or of worshipful feelings. Worshipful feelings may be involved in 1 (c) in some contexts at some level, and are at their highest in 43, or 98, but they do not override the woman's inadequacy (in 43, or 98, the very context of situation makes us aware of Natasha's guilt as she comes to ask the dying Bolkonsky for forgiveness). In the contexts where wonder and admiration are at the forefront, there is no mention of a woman. Considering that Pushkin's stanza is not addressed to a divinity but to an attractive woman (a genius of true beauty), there is strong evidence suggesting that there is something wrong with the woman and/or her appearance. In other words, 'before' and 'appear' create a situational semantic prosody, similar to Sinclair's 'when she was' (2006), which is singularly fractured in Pushkin's stanza. As it is not ironic [we have seen instances of irony created by these collocates in 4 (c)], it must fall into the 'insincerity' category, possibly into its 'wishful thinking' subgroup (Milojkovic 2013a; see also Section 9).

As for the relationship of power that is often foregrounded in the states of affairs created by these two collocates, one party is often inferior to the other. In Pushkin's poem,

ostensibly the woman is superior, while the states of affairs created by ‘before’ and ‘appear’, as well as the subtext (so far based on 5 lines only, but to be studied in more detail further in the text) claim that she is deficient. Which one is the meaning? Can both be the truth perceived by the poet, one attitude overriding the other, also present? This would seem to be the most probable answer.

It is important to note that there are 3 instances of Pushkin’s co-selecting these collocates in the whole of his authorial corpus. All 3 are among these findings, and all 3 involve some sort of deception:

- Context 6 from 1 (c) (the heroine disguises herself so as not to be recognized by her lover),
- Context 5 from 2 (c) (the young man adopting the then fashionable pose of disappointment, which captivates all the girls in the neighbourhood), and
- Context 93 from 4 (c) (the soldiers encounter a sudden military threat when reaching a deceptively peaceful meadow).

In 2 contexts out of the 3 (contexts 6 and 5) the deception is deliberate. It is remarkable that both describe an unnatural pose, greatly deviating from reality, in the context of man-woman relationships. These 2 contexts are from the same novel (*The Squire’s Daughter,* 1831). They must be linked to the stanza by Zhukovsky, which triggered the exploration of the states of affairs chunked by ‘before’ and ‘appear’ in the RNC. That stanza is sinister – it describes the beautiful countess masked as a dead man.

All in all, co-selecting *перед* (‘before’) and *явиться* (‘appear’) in the RNC yielded the following relevant data:

- in the context similar to Pushkin's context of situation, where a woman appears before a man who finds her attractive, there is a strong sense of the woman’s inadequacy or guilt.
- in the contexts where wonder and admiration are at the forefront, there is no

mention of a woman.

- all 3 contexts authored by Pushkin involve deceptive qualities; 2 of them involve man-woman relationships, and disguises and/or unnatural poses.

As the co-selection process was one-way only (it only took into account the contexts where the preposition preceded the verb, as in Pushkin's original), these states of affairs must have a bearing on the final verdict regarding subtext in the line under investigation.

5.4.3 Researching a shorter grammatical string for quasi-propositional variables

In order to investigate subtext in Pushkin further – 5 lines in Figure 5.5 is nowhere near enough – I will research a shorter string for states of affairs it creates. My shorter string will be **ась перед* ('*ed before'): the verb ending in Russian denotes the feminine gender, the personal pronoun in the second person singular *ты* ('thou'), found in Pushkin, is here omitted.

The main corpus of the RNC yielded 1404 contexts. I studied the first random 200 for quasi-propositional variables and the contexts themselves. Firstly, I found it necessary to group the verbs thematically, as the contexts seemed ready to be grouped. It will prove useful when I need to justify my decisions and judgements as to what the QPVs actually represent.

This is the thematic grouping of the QPVs, regardless of whether they appear once, twice, or many times. Not all these contexts of situation feature a woman, as gender is a grammatical category in Russian. If a woman is involved, it will be stated in the description of the context.

1. QPVs in contexts describing an image unfolding in reality or in one's mind's eye (36):

- *открылась* ('opened', perfective aspect) 4
- *открывалась* ('opened', imperfective aspect,) 2
- *раскрылась* ('was exposed', perfective aspect) 4

- *раскрывалась* (‘was exposed’, imperfective aspect)
- *разверзлась* (‘yawned’, ‘gaped’, perfective aspect) 2
- *разверзалась* (ibid., imperfective aspect)
- *разворачивалась* (‘unfolded’, imperfective aspect) 3
- *развернулась* (ibid., perfective aspect)
- *развертывалась* (‘unrolled’, imperfective aspect)
- *пронеслась* (‘rushed’, perfective aspect) 2
- *распахивалась* (‘flung open’, imperfective aspect)
- *распахнулась* (ibid., perfective aspect)
- *простиралась* (‘stretched’, imperfective aspect) 2
- *распростерлась* (ibid, perfective aspect)
- *растекалась* (‘spread’, imperfective aspect)
- *нарисовалась* (‘drew itself’, perfective aspect)
- *изобразилась* (‘painted itself’, perfective aspect)
- *расстилалась* (‘lay’, imperfective aspect)
- *удлинялась* (‘became longer’, imperfective aspect)
- *обнажилась* (‘lay bare’, perfective aspect)
- *поднялась* (‘rose’, perfective aspect)
- *высветилась* (‘zoomed in’, perfective aspect)
- *складывалась* (‘put together’, imperfective aspect)
- *повторялась* (‘recurred’, imperfective aspect)

2. QPVs in contexts where the QPV means ‘stopped’:

- *остановилась* (perfective aspect) 27
- *останавливалась* (imperfective aspect) 2

3. QPVs in contexts where the QPV means ‘apologised’:

- *извинилась* (perfective aspect) 6
- *извинялась* (imperfective aspect) 4

4. QPVs in contexts where the QPV means ‘appeared’:

- *появилась* (perfective aspect) 7

- *явилась* (perfective aspect, as in Pushkin)
- *являлась* (imperfective aspect)

5. QVPs in contexts referring to a crowd, a queue etc:

- *выстраивалась* ('formed', imperfective aspect)
- *выстроилась* ('formed', perfective aspect)
- *собралась* ('gathered', perfective aspect) 2
- *расступилась* ('parted' (to let someone pass), perfective aspect)
- *расступалась* (ibid., imperfective aspect)
- *разбивалась* ('broke', imperfective aspect)
- *раздалась* ('made way', perfective aspect)

6. QPVs in contexts describing a woman before a mirror:

a) in the meaning of 'turn around', 'spin', 'whirl'

- *вертелась* (imperfective aspect) 4
- *повертелась* (perfective aspect) 2
- *завертелась* ('started spinning', perfective aspect)
- *крутилась* (imperfective aspect) 2
- *покружилась* (perfective aspect)

b) in the meaning of 'enjoyed her reflection':

- *охорашивалась* ('preened herself', imperfective aspect)
- *упивалась* ('feasted her eyes', imperfective aspect)

c) other meanings

- *смотрелась* ('looked at herself', imperfective aspect)
- *причесывалась* ('combed her hair', imperfective aspect)
- *разделась* ('undressed', perfective aspect)

7. Contexts describing a woman kneeling:

- *опустилась (на колени)* ('went down (on her knees)), perfective aspect 3
- *бросилась (на колени)* ('threw herself (on her knees)', perfective aspect)
- *исползалась на карачках* ('crawled on all fours', perfective aspect)
- *становилась (на одно колено)* ('went down on one knee', imperfective aspect)

8. Contexts describing a woman forfeiting control:

- *растерялась* (‘was taken aback’, perfective aspect) 2
- *терялась* (‘became flustered’, imperfective aspect)
- *склонилась* (‘bowed’, ‘bent’, perfective aspect) 2
- *склонялась* (ibid., imperfective aspect)

9. Contexts where the QPV means ‘justified herself to someone’, with a woman agent

- *оправдалась* (‘justified herself’, perfective aspect)
- *оправдывалась* (ibid., imperfective aspect) 2
- *пыталась оправдываться* (‘attempted to justify herself’, imperfective aspect)

10. Contexts where the QPV means ‘report to’, with a woman agent:

- *отчитывалась* (imperfective aspect)
- *отчиталась* (perfective aspect)

11. Contexts where the QPV means ‘grovel’, with a woman agent:

- *пресмыкалась* (imperfective aspect)
- *стелилась* (imperfective aspect)

12. contexts where the ‘woman’ wants to make herself attractive or available to a man (the ‘woman’ is in quote marks because in one context she appears in a dream, and in another it is a bomb in a dream, acting like a woman):

- *разделась* (‘undressed herself’, perfective aspect) 2
- *раздевалась* (ibid., imperfective aspect)
- *прошлась* (‘walked’, perfective aspect)
- *вытягивалась* (‘stretched’, imperfective aspect)
- *охорашивалась* (‘preened herself’, imperfective aspect)

13. contexts of the woman’s anxiety:

- *волновалась* (‘worried’, imperfective aspect) 3
- *всполошилась* (‘took alarm’, perfective aspect)

14. unclassified (these contexts did not appear ‘willing’ to belong to a group)

The contexts above were not grouped according to the context of situation, but

rather according to the meanings of the verbs. Nevertheless, many contexts analysed here involve the woman as a relevant person, in situations which are reminiscent of some of the 5 contexts found in the initial concordance (Figure 5.5). The final analysis of these QPVs will be postponed until Louw's procedure has been followed. Unlike the grouping above, the following statistics is founded on Louw's understanding of subtext: the QPVs are presented in order of frequency, notwithstanding the context of situation or the meaning of the verbs in their respective contexts. Those verbs which occur only once were not included, unless they are a different (perfective or imperfective) aspect of another appearing verb. The English equivalent is given first:

1. 'stopped': *остановилась 26 останавливалась 3*
2. 'apologised': *извинилась 6 извинялась 4*
3. 'appeared': *появилась 7 явилась 1 являлась 1*
4. 'opened': *открылась 5 открывалась 2*
5. 'turned around': *вертелась 4 повертелась 2 завертелась 1*
6. 'found herself/itself': *оказалась 5*
7. 'was exposed': *раскрылась 4 раскрывалась 1*
8. 'unfolded': *разворачивалась 3 развернулась 1*
9. 'lowered herself': *опустилась 3*
10. 'worried': *волновалась 3*
11. 'rushed': *пронеслась 3*
12. 'undressed': *разделась 3*
13. 'sat down': *уселась 3*
14. 'bowed, bent': *склонилась 2 склонялась 1*
15. 'justified herself': *оправдывалась 2 оправдалась 1*
16. 'gaped', yawned': *разверзлась 2 разверзалась 1*
17. 'stretched (of a landscape)': *простиралась 2*
18. 'turned around': *крутилась 2*
19. 'preened herself': *охорашивалась 2*

20. ‘found itself’, ‘was situated’: *находилась 2*
21. ‘found herself’ *очутилась 2*
22. ‘remained’: *осталась 2*
23. ‘was taken aback’: *растерялась 2*
24. ‘gathered’: *собралась 2*
25. ‘reported to’: *отчитывалась, отчиталась*
26. ‘formed (of a queue)’: *выстраивалась, выстроилась*
27. ‘parted (e.g. a crowd)’: *расступилась, расступалась*

The thematic grouping of contexts containing the QVPs, which preceded this impartial classification, was helpful because the researcher was in no danger of letting a verb relexicalise outside its context. Now we know that all verbs describing turning around either involve a mirror or a male observer; all verbs translated as ‘lowered herself’ involve kneeling; all instances of ‘undressed’ take place either before a man (2 times) or a mirror (once); all instances of ‘opened’ involve landscapes or mental images, and not doors, for example. If the thematic grouping had been omitted, a mere list of the QPVs would not be sufficiently informative. The knowledge of the QVPs’ contextual meanings allows the following classification, according to the trends of meaning and in order of frequency:

a) the meaning of ‘stopped’, 29 verbs: *остановилась* 26 *останавливалась* 3
b) the meaning of ‘inferiority or absence of control’, 28 verbs:

- ‘apologised’: *извинилась 6 извинялась 4*
- ‘lowered herself’: *опустилась 3*
- ‘worried’: *волновалась 3*
- ‘bowed, bent’: *склонилась 2 склонялась 1*
- ‘justified herself’: *оправдывалась 2 оправдалась1*
- ‘was taken aback’: *растерялась 2*
- ‘reported to’: *отчитывалась 1, отчиталась 1*
- ‘parted (e.g. a crowd)’: *расступилась 1, расступалась 1*

c) verbs that are used to describe a physical image, like a landscape, or mental image, apart from the general ‘appear’, 25 verbs:

- ‘opened’: *открылась 5 открывалась 2*
- ‘was exposed’: *раскрылась 4 раскрывалась 1*
- ‘unfolded’: *разворачивалась 3 развернулась 1*
- ‘rushed’: *пронеслась 3*
- ‘gaped’, yawned’: *разверзлась 2 разверзалась 1*
- ‘stretched (of a landscape)’: *простиралась 2*

d) verbs sharing aspects of the meaning of the verb ‘to be’, like ‘appeared’, 20 verbs:

- ‘appeared’: *появилась 7 явилась 1 являлась 1*
- ‘found herself/itself’: *оказалась 5*
- ‘found itself’, ‘was situated’: *находилась 2*
- ‘found herself’ *очутилась 2*
- ‘remained’: *осталась 2*

e) implications of sexual attractiveness in appearing before men or mirrors, 14 verbs:

- ‘turned around’: *вертелась 4 повертелась 2 завертелась 1*
- ‘undressed’: *разделась 3*
- ‘turned around’: *крутилась 2*
- ‘preened herself’: *охорашивалась 2*

f) verbs referring to queues or crowds, 4 verbs:

- ‘gathered’ (of a crowd): *собралась 2*
- ‘formed’ (of a queue): *выстраивалась 1 , выстроилась 1*

g) ‘sat down’, with a woman agent, 3 verbs (one of these occurrences also involved a mirror)

The categories (a), (b), (c), (d), and (e), their order determined by their frequency, roughly correspond to the 5 most frequent QVPs:

1. ‘stopped’: *остановилась 25 останавливалась 3*
2. ‘apologised’: *извинилась 6 извинялась 4*

3. ‘appeared’: *появилась 7 явилась 1 являлась 1*
4. ‘opened’: *открылась 5 открывалась 2*
5. ‘turned around’: *вертелась 4 повертелась 2 завертелась 1*

So far it appears that the two trends of meaning seen in the 5-line concordance (Figure 5.5) – sexual provocation in 1 line and guilt in 3 lines – are supported by categories (e) and (b), respectively, of the classification given above. The proportion of these QPVs is 1 to 3 in the concordance of *ты *сь перед** (‘you (thou) *ed before’), and it is 14 to 28 in **ась перед*. But what do we make of the most frequent QPV, ‘stopped’?

5.4.4 The implications of the most frequent QPVs in **ась перед*

When looking through the group of verbs whose meaning may be translated as ‘unfolded’ or ‘opened before’, one is struck by the diversity of the Russian verbs: so many different verb prefixes, suffixes and roots, so many nuances of meaning to describe a more or less similar phenomenon of a picture or an image unfolding or passing before one. According to Louw’s principle, this very diversity prevented these verbs to be strictly considered quasi-propositional variables, due to the low frequency of each – the majority occurred only once in 200 contexts. It stands to reason that the most frequent QPV is a very ordinary and thus a not very ‘telling’ verb: *остановилась* (‘stopped’ in the perfective aspect). Although the gender of the verb is feminine, the agent does not have to be a woman: it can be any feminine noun, e.g. a car (*машина*), or a group of schoolchildren (*группа школьников*), or an abstract quality expressed by a feminine noun. According to Louw, the most frequent QVPs are the subtext of a grammatical string, its hidden meaning. What hidden meaning to the words of adoring Pushkin can be provided by the ordinary action of stopping? There is no obvious guilt, no unequivocal sexual provocation, we stop every day and many times a day.

Ordinary verbs can have their special roles in literary texts – and the texts provided by the RNC, at least on this occasion, are mostly literary. Any text is written for a purpose, to tell us what we do not know, to describe events leading to a conclusion that is important

to the writer. A writer is not likely to tell us that after a person stopped, another ordinary and unremarkable event took place. In all probability, after someone or something stops in a literary world, something else, quite relevant, is about to start.

Since the verb 'to stop' is in itself ordinary rather than telling, let us see what transitions it facilitates in its contexts. In its imperfective aspect it appears only 3 times. One of the contexts is the criticism of 'the circles of liberal opposition' which 'did not stop before the danger of changing horses while crossing the river':

Эти списки возникали в кругах либеральной оппозиции, которая не останавливалась перед опасностью «перемены коней на переправе […] [Г. Иоффе. Русский либерал. Премьер-министр Временного правительства — князь Львов // «Наука и жизнь», 2006]

(These lists were created in the circles of liberal opposition, which did not stop before the danger of changing horses while crossing the river […] [G. Ioffe. The Russian Liberal. Prime Minister of Russian Provisional Government – Prince Lyvov// 'Scinece and Life', 2006])

The wider context is not overtly critical, but 'danger' is the preposition's closest collocate, and the action takes place during the turbulent times preceding the Russian Revolution. In another context with the same verb a young woman 'stopped before the door leading to the apartment of her childhood' with a mixed feeling, in which 2 ingredients out of the 4 are negative:

Каждый раз, когда Женя останавливалась перед дверью квартиры, где прошло её детство, она испытывала сложнейшее чувство: умиление, гнев, тоску и нежность. [Людмила Улицкая. Казус Кукоцкого [Путешествие в седьмую сторону света] // «Новый Мир», 2000]

(Every time when Zhenya stopped before the door to the apartment where she had spent her childhood, she experienced the most complex feeling: tenderness, wrath, melancholy, and affection.[Lyudmila Ulitskaya. The Incident of Kukotsky. [the Journey to the Seventh Quarter]//Novy Mir, 2000])

Along with affection and tenderness the heroine experiences wrath and melancholy. I will put down both these contexts as those of critical transition with a difficult past and an

uncertain future (both wider contexts warrant it too). The third context clearly expresses a negative attitude to what is described after 'stopped before':

Мог быть груб и темен мужик, глупа и вздорна баба, но даже их невежественность останавливалась перед оскорблением убогого, осквернением красивого, восхищением насилием, оправданием подлости, предательства, корысти [...][«Futurum», или О разделении России на «мы» и «они» (2004) // «Жизнь национальностей», 2004.06.16]

(No matter how coarse and benighted was the peasant, how foolish and shallow the peasant woman, but even their ignorance stopped before [short of] offending the crippled, profaning the beautiful, delighting in violence, excusing meanness, treachery and self interest [...] ['Futurum', or On the division of Russia into 'us' and 'them' (2004) // 'Life of the Nationalities', 16.06.2004])

Among the 3 contexts of *останавливалась перед* 2 are political and one is personal. All 3 clearly refer to a possible transition. The transition is definitely harmful in the political contexts, and inspires certain negative feelings in the context depicting the young woman (even though wrath and melancholy are accompanied by tenderness and affection, one's childhood memories should ideally be more positive). The general situational prosody of *останавливалась перед* seems to be that of a negative prosody of transition.

In my dealing with the contexts of the same verb ('stopped') in the perfective aspect – *остановилась*, I will adopt the same method and look at not only close collocates but also words that appear farther away from the studied string. If I do not get enough information from the immediate collocates, I will look at a still wider context. This will ensure I understand the nature of the transition expressed by *остановилась*.

Out of the 25 contexts, 17 (68%) were negative in varying degrees, and one I pronounced neutral. Out of the remaining 8, 3 were very positive, and 5 were ironic. The positive usages appeared in contexts of situation very different from the one discussed in Pushkin: a woman admiring a view in 2 cases, and a waitress bringing a drink. Speaking of the context of situation, among the 17 contexts which I found negative, in 8 the subject of the verb 'stop' was a woman:

- In 3 contexts she ‘stopped before’ a man (relevant personality) during a dispute growing into a quarrel (all 3 contexts are by the same author).
- In 2 contexts, the woman stops before a door or a building where the man, the ‘relevant personality’, lives. In one context she dislikes the man, in the other loves him to distraction, but in both cases the heroine feels that a question of life and death may be solved inside.
- In 1 context, *остановилась* is a transition from a routine crossing the street to watching a passer-by suffer a heart attack.
- 2 contexts describe a humbling experience: one is of a woman before a woman (during a greeting); the other is of a man before a woman (during an apology). The latter relationship is that of love. *Остановилась* is a transition before the other, dominant party makes the next move

Although no context, situationally, perfectly resembles Pushkin’s, these findings point to the conclusion that the semantic prosody of the sequence *остановилась/ останавливалась* перед (‘stopped before’) is that of a (potentially) harmful transition. In the whole of 28 contexts, there are 5 exceptions to this prosody, 3 of which involve women (not counting 5 ironic usages). In 2 contexts a woman is admiring a view, in one she is receiving her lover’s apology after a quarrel. On the basis of these findings we may conclude that the sequence in question mostly prospects a potentially dangerous transition, unless the agent is a woman admiring a view.

Let us now consider the other ‘less telling’ group of QPVs which stands out in the classification because it contains the verb actually used by the poet. It is group (d), consisting of 20 QPVs which share the semantic feature of the verb ‘to be’:

- ‘appeared’: *появилась 7 явилась 1 являлась 1*
- ‘found herself/itself’: *оказалась 5*
- ‘was situated’: *находилась 2*
- ‘found herself’: *очутилась 2*
- ‘remained’: *осталась 2*

These are the features of these QPVs in contexts:

1. The contexts of the last 4 QPVs - *оказалась* (literally 'found herself', grammatical feminine gender), *находилась* (literally 'was situated'); *очутилась* (literally 'found herself') ; *осталась* ('remained') - contain negative collocates and/or describe challenging transitions.

2. The context of *являлась*'(the imperfective aspect of Pushkin's 'appear'), happily, comes from a text by Pushkin, and hence is of interest. (I discuss his rare usages of the perfective aspect of the verb, apart from the studied stanza, in Table 5.4– one in the context of military danger and 2 in the context of poses and disguises.) The context discusses the actress Semenova, much appreciated by Pushkin.[21] Pushkin criticises her partner in the following sentence: '*Долго Семёнова являлась перед нами с диким, но пламенным Яковлевым, который, когда не был пьян, напоминал нам пьяного Тальма.'* (For a long time did Semenova appear before us with the wild but ardent Yakovlyev, whose acting, when he was not drunk, was similar to that of a drunken Talma.) Pushkin's critical tone is not directed against the actress but against her partner; however, *являлась* here does prospect negativity.

3. The context of *явилась* is ironic and does not involve a woman.

4. The context of *появлялась* (imperfective aspect) is the appearance of a famous ballerina, 'small, with a queenly erect back and a quiet domineering voice'.

5. *Появилась* (perfective aspect) was present in 7 contexts. In one context the subject was inanimate; the rest featured

- an actress.
- a female demon.
- a woman hit by a car.
- a woman calculating to impress.
- a woman who is attractive but troubled

[21] Pushkin wrote: 'Семенова не имеет соперницы... она осталась единодержавною царицею трагической сцены.' (Semenova has no rival… She has remained the sovereign empress of the tragic stage.)

It seems that all these verbs tend to appear in negative or challenging contexts, whether they involve women as participants or not. In particular, if a woman participant is spared censure or misfortune, she happens to be a ballerina or an actress. This is in accordance with Pushkin's usage of *явилась* (the same verb but perfective aspect), co-selected with *перед* ('before') in the RNC, which, along with the context of military danger, yielded 2 contexts of poses and disguises as part of the strategy of romance. Poses and disguises may involve as much acting as a stage performance.

All in all, the search for the quasi-propositional variables of the string **ась перед* ('*ed before'), which is contained in the string *ты *сь перед* ('you *ed before', see also Figure 5.5), yielded several groups of quasi-propositional variables. Their implications will be summarized in the paragraphs below. It must be repeated at this point that the string studied initially, *ты *сь перед* ('you *ed before', Figure 5.5), is not the actual grammatical string used in Pushkin's stanza. It is the string used in modern Russian, whereas Pushkin's word order is reversed. The decision to change the word order was resorted to, because there are no occurrences of Pushkin's grammatical string in the corpus, except as a paraphrase of the verse under discussion. The modern grammatical string yielded only 5 occurrences (Figure 5.5), hence the need to study the shorter string contained in it. The search yielded several groups of verbs. From each group, the most frequent verbs were:

1. 'stopped': *остановилась 26 останавливалась 3*
2. 'apologised': *извинилась 6 извинялась 4*
3. 'appeared': *появилась 7 явилась 1 являлась 1*
4. 'opened': *открылась 5 открывалась 2*
5. 'turned around': *вертелась 4 повертелась 2 завертелась 1*

Since the QPVs had been previously grouped thematically according to their contexts, the situational meaning of categories 2, 4, and 5 was clear. It was obvious that categories 2 and 5 corresponded to 4 lines of the 5-line concordance of *ты *сь перед** ('you *ed before', Figure 5.5). Category 2 corresponded to lines 3, 4 and 5 in the

concordance of *ты *сь перед**, pointing to the woman's guilt. Category 5 corresponded to line 2 of the concordance in Figure 5.5 (a woman strutting attractively before a man), and possibly even to the trivial context in line 1 (a woman washing her face before applying a cream), since many QPVs from Category 5 come from contexts involving a mirror.

These findings are encouraging, as the subtext of the string *ты *сь перед** ('you *ed before') has obviously been confirmed by the subtext of the shorter string, **ась перед* ('*ed before'), contained in it. As for the other 3 categories of QPVs, category 4 referred to landscapes and open spaces, and need not be considered as subtextual in Pushkin's case. The 2 remaining ones, 1 and 3, were studied further because they were, in terms of meaning, too ordinary to reveal any information. The imperfective aspect of the Russian equivalent of 'stopped before' turned out to contain a negative prosody of transition; furthermore, 2 contexts out of 3 pointed to danger and harm. The perfective aspect of the same verb appeared in predominantly negative contexts (68%). If the agent was a woman, the prosody was that of a potentially dangerous transition, unless the woman is admiring a view.

Category 3 described either a challenging transition or the appearance of an attractive woman displaying a negative quality; the woman was not disapproved of if she was an actress. This is an important finding since this category is semantically closest to Pushkin's lexical choice. Crucially, we had already discovered 3 contexts authored by him, where he co-selects the verb and the preposition used in the stanza to create contexts of poses, disguises, or the theatre, as shown in the next section.

5.4.5 Pushkin's use of the verb *являлась*

It is interesting that among the contexts of **ась перед* there was one context by Pushkin containing *являлась* ('appeared', imperfective aspect), which I found positive as to the verb's subject but critical as to the the context's message. This made me search for the verb *являлась* in Pushkin's subcorpus provided by the RNC. Even though in the poem studied the aspect is perfective, both these verbs take the same feminine ending in the past. 3 contexts were found, among which was the one already quoted:

1 *Долго Семёнова являлась перед нами с диким, но пламенным Яковлевым, который, когда не был пьян, напоминал нам пьяного Тальма. [А. С. Пушкин. Мои замечания об русском театре* (1820)]

(For a long time did Semenova appear before us with the wild but ardent Yakovlyev, whose acting, when he was not drunk, was similar to that of a drunken Talma. [A. S. Pushkin. *My remarks on the Russian theatre*. (1820)])

2 *Но душный воздух и закрытые окна так мне надоели во время болезни моей, что весна являлась моему воображению со всею поэтической своей прелестию. [А. С. Пушкин. Записные книжки (1815-1836)]*

(But the stuffy air and the closed windows bored me so much during my sickness that spring appeared to my imagination with all its poetic charm. [A. S. Pushkin. *Notebooks*. (1815-1836)])

3 *Редко наша красавица являлась посреди гостей, пирующих у Кирила Петровича. [А. С. Пушкин. Дубровский (1833)*

(Rarely did our beauty appear among the guests feasting at Kirila Petrovich's. [A. S. Pushkin. *Dubrovsky* (1833)]

Figure 5.7 The contexts of the verb form *являлась* in Pushkin's subcorpus (RNC)

In Context 1, the preposition 'before' is employed, and the woman is an actress. In Context 3, the preposition is 'among'. In Context 2, the male persona is mentioned using the dative case, and this positive mention of spring is in accordance with the data from the Poetry Corpus in Section 5.2, where visions appeared before male participants who were predominantly referred to in the dative case.

Although the already quoted context featuring actress Semenova is the only one where Pushkin uses the preposition *перед*, in the 2 other contexts the verb's subjects are also very much approved of by the author. In the second context it is personified spring which appears to the poet's imagination 'with all of its poetic charm'. In the third context the beautiful heroine of Pushkin's novel *Dubrovsky*, who has the author's every sympathy, is said to appear only rarely among her father's feasting guests.

Surprisingly, all 3 contexts share a particular quality: the subject of the verb has the author's complete admiration, but she appears in a context that the author finds negative. It

must be concluded that Pushkin uses this form of the verb to contrast its subject, which he admires, with the flawed surroundings – a bad co-actor, a long illness, a drunken feast – whether he does use the preposition *перед* (‘before’) or not.

5.4.6 Establishing the subtext of *передо мной явилась ты* (‘you appeared before me’)

It is now my aim to establish whether and to what extent the 5 subtextual lines from the RNC may be deemed reliable in determining the subtext of Pushkin’s line *Передо мной явилась ты*. I have carried out several corpus searches and need to generalize on their findings. Given the diversity of these, it will not be easy.

The starting point was the RNC search for *ты+*сь+перед** in the RNC, which gave 5 lines overall (Figure 5.5). Of these, 1 QPV was trivial, 1 suggested sexual provocation on the part of the woman addressed in the line, and 3 pointed to the woman’s guilt.

First, I decided to support my claim that the lexical item in Pushkin’s line, the verb *явилась* (‘appeared’), is relexicalised by *мимолетное виденье* (‘transient vision’) in the next line. I found 8 contexts in the Poetry Corpus of the RNC to support this claim, while 3 contexts contained a delexical mention of a beautiful woman as a vision (Figure 5.6).

Pushkin’s stanzas aside, out of the 11 texts where *явиться* (‘appear’) collocated with *видение* (‘vision’), only one, written by Pushkin’s contemporary Zhukovsky, featured the preposition *перед*. This was a sinister context involving a young and beautiful woman wearing a dead man’s mask. The context of situation was very similar to Pushkin’s – a brief appearance of an attractive woman before a man. The other contexts, if they mentioned a person witnessing an apparition, featured the dative case and not the preposition *перед*.

This made me co-select the preposition *перед* (‘before’) and the verb *явиться* (‘appear’) in the main corpus of the RNC, in this order, as I suspected they might create negative states of affairs. The main corpus of the RNC yielded 127 contexts. I ignored those where *перед* (‘before’) had a temporal meaning, and classified the remaining ones

into subgroups (Table 5.4). As regards the context where a woman appears before a man, her attractiveness was a relevant feature, but there was something odd either about her or the fact of her appearance. The most positive context described Natasha Rostova's appearance before mortally wounded Bolkonsky in *War and Peace* by Tolstoy – she has all the sympathy of the author, but has come to ask forgiveness for her past betrayal. This seemed to correspond to Pushkin's subtext as shown by the 5-line concordance (Figure 5.5) – the sexual attraction is mentioned in the Tolstoy excerpt (*та, которую он желал* – 'she whom he desired'), and the guilt of betrayal and the willing forgiveness permeates their meeting. The woman generally controls the situation. In contexts involving the supernatural, apart from the appearance of Christ and saints, ghosts were harmless, apparitions positive, and creatures from hell punished almost exclusively sinners. In contexts termed miscellaneous, 61% of contexts were negative, and the few positive ones involved the appearance of divinity or male heroes. Thus, worshipful attitudes never involved women. Incidentally, 3 contexts came from Pushkin himself – 2 involved poses (Table 5.4, 2 (c), context 5) and disguises (Table 5.4, 1 (c), context 6) as part of romantic play, and one an unexpected military danger (Table 5.4, 4 (c), context 93). What the 3 contexts by Pushkin shared was the feeling of deceptive appearance, not viewed negatively in romantic situations.

Finally, I went back to subtext and searched the line **ась перед*, which was contained in the first studied subtextual string (Figure 5.5). Here it is worth repeating that the word order in the studied string was not the same as in Pushkin, because Pushkin's is poetic, archaic and does not exist in any corpus except as a quote from Pushkin's poem. This gives additional importance to coselection-based searches in this section. The QPVs in the shorter string corresponded to those of *ты+*сь+перед** in that there were 2 large groups of verbs: those implying loss of face or control (28 occurrences) and those implying overt implication (of the importance) of sexual attractiveness (14 verbs). There were also 29 verbs with the meaning of 'stopped' and 20 verbs sharing meaning with 'appear', 9 of which corresponding to Pushkin's chosen verb. Out of the contexts of 'stopped', only 3 were positive and they did not involve women appearing before men. All the verbs grouped

semantically with ‘appear’ were either negative or described challenging transitions. As for the verbs corresponding to ‘appear’ in meaning (9 verbs), 3 described the appearing women favourably (as free of censure or misfortune); 2 of these women were actresses and one was a very famous ballerina, queenly in appearance. One of the contexts, featuring the verb *являлась,* the imperfective variant of *явилась,* belonged to Pushkin. It described an actress very much admired by him (Semenova) in the company of an actor whom he detested. By searching for *являлась* in Pushkin’s authorial corpus, I discovered 2 more contexts featuring this form but not co-selected with *перед* (‘before’). Both of these described, as in the Semenova context, the appearance of something or someone very much admired and very beautiful (namely, spring and the heroine of *Dubrovsky*) in a very negative context (Figure 5.7).

Although the subtext of the line *ты+*сь+перед** (Figure 5.5) corresponds to that of **ась перед*, this finding is not final in the case of Pushkin’s stanza in question because the word order of these strings is not Pushkin’s. The former string was researched because its word order is actually used in (modern) Russian. But in the light of the fact that Pushkin’s word order places the preposition at the beginning of the string, the co-selected *перед* (‘before’) and *явиться* (‘appear’) in this order (Table 5.4) yield states of affairs that might be scientifically even more relevant. These states of affairs, if they feature a woman appearing before a man, point to the woman’s inadequacy or guilt. This corresponds even to Zhukovsky’s line – the beautiful countess giving the poet a sudden shock was wearing a mask. Zhukovsky was Pushkin’s contemporary and his stanza was the only instance in the whole of RNC where *явиться* (‘appear’) and *видение* (‘vision’) were co-selected, in this order, in the presence of the preposition *перед* (‘before’). Pushkin, however, co-selects *перед* (‘before’) and *явиться* (‘appear’) in a prose piece when describing harmless poses and disguises.

All in all, the subtext of the line *ты+*сь+перед** (Figure 5.5) in Russian tends to speak negatively of the attractive woman, while that of **ась перед* supports the former but adds the element of a potentially dangerous transition – bar those situations where the woman is an actress, a performer. This must be linked with poses and disguises in Pushkin

and Zhukovsky, and with Pushkin's use of the imperfective equivalent of *явилась*'– *являлась* (Figure 5.7). The latter is used by Pushkin 3 times when he describes an object of admiration as well as of great beauty – and, by the way, not an intimate companion.

Pushkin's individual use suggests that he viewed his heroine as an appearance of great beauty – the genius of true beauty – rather than as a 'person, personality'. He may have had a similar attitude to women in general, or to this particular woman at the moment of writing. Strictly speaking, the longer the string, the more precise the subtext – therefore, scientifically, authorial meaning cannot be fully ascertained until the corpus grows big enough. But the context of situation claiming that the woman appearing before you on the stage may be viewed positively and with admiration, combined with Pushkin's customary use of the verb, leaves us with the statement that he does admire her for her great beauty – as he tells us himself. However, the woman in question is not an actress, or a book character, or spring. Perhaps Pushkin courteously creates, by the collocates 'appeared' and 'fleeting vision', and by the admiring tone, a state of affairs which goes well with an actress appearing as part of her profession, at the same time as universal subtext whispers the outwardly blocked censure of the personality of the apparition.

5.5 Literary device and authorial intention in Pushkin's *передо мной явилась ты* ('you appeared before me')

The preceding sections prove successfully that the subtext in Pushkin betrays a negative aspect not ostensibly present in the poem. On the other hand, what is ostensibly present is the literary device:

Передо мной ***явилась*** *ты,*	You **appeared** before me
Как ***мимолетное виденье****…*	Like a **fleeting vision**…

'Like a fleeting vision' is a simile – we know this because of the presence of 'like'. It is directly dependent on the verb in our target grammar string, which makes the verb 'appeared' part of the device. The presence of 'like' is supported by relexicalisation,

confirmed in Section 5.4.2. However, the compliment paid by the device is sustained only at the level of lexical collocation. The appearance of visions in the Poetry Corpus of the RNC is expressed by using the dative case when referring to the male witness, not the preposition *перед* ('before'). This preposition, a grammar word, is what undermines the lexical in Pushkin, because, when co-selected with the verb, it prospects negativity in the RNC contexts of situation where women appear before men (while visions and apparitions as such predominantly have a positive role in the RNC contexts). The negativity prospected by this preposition also applies to its grammar string, with the verb wildcarded. This instance of grammar overruling the implications of purely lexical co-selection supports Louw's claim that logic comes before metaphysics (Louw and Milojkovic 2016: 189-190). This is the most important finding of this section.

If we add Pushkin's line to the English examples studied in Section 5.3, this is what Table 5.3 looks like:

Table 5.5 Table 5.3 ('Literary devices and authorial intention established by means of relexicalisation and context clues') revised

text/ lexical choice, subtext and context clues	1. Is the author's lexical choice relexicalised?	2. Do the context clues support the subtext of the grammatical string?	3. Is the lexico-grammatical collocation a device?	4. Does the lexico-grammatical collocation express authorial intention?
'that is no *country* for'	no, it is used in its full lexical meaning	no, the context clues do not support 'reason' or 'excuse'	no	no
'(a heart) made *bleak* by'	poorly	yes	no	yes
'now that my *ladder*'s gone'	yes	yes	yes	yes
'you *appear*ed before me'	yes	no	only at the level of the lexical	no

Section 5.3 of this section claimed that 'for a lexico-grammatical collocation to be

pronounced a device, questions 1 and 2 in Table 5.3 would be answered affirmatively. For establishing authorial intention, it is enough to find that the author's subtext is supported by context clues – therefore, it is enough to answer question 2 in Table 5.3 affirmatively.' How are we to pass judgement on devices which are sustained at the level of the lexical, but flouted by their own grammar? In my view, they ought to count as devices because they reflect authorial intention, as Pushkin's lexis undoubtedly does. This opens the question of the depth of devices: while Yeats's device runs deep, Pushkin's is superficial and undermined by contradicting subtext. These subtleties must be felt by native speakers of these languages and must account for the apparently inexplicable poetic appeal of the studied lines.

5.6 Implications for existing scholarship on authorial intention

This is the first instance of authorial intention arrived at by corpus stylistic means (Wang and Humblé 2017: 551). It would seem suitable at this point to contextualise this approach within the general academic thought. Sotirova (2015: 132-148) provides a historical overview of how authorial intention was seen to participate in the critical appreciation of a poetic text. According to Plato and Socrates, the poet might not be aware of the full meaning of his words. For Aristotle, the excellence of a rhetorical argument depends on how well the author's intention is conveyed, but this does not apply to poetry (Halliwell 1986: 191, quoted on p. 133). For the medieval tradition, authorial intent is essential to meaning, whereas Montaigne believes the author to be capable of producing a richer meaning than intended. Wimsatt and Beardsley (1954: 3-11, quoted on p. 134) quarrel with the 'romantic intentionalism', described by Mitscherling et al. (2004: 39, quoted on p. 135) as the author experiencing a psychic state, and the reader then re-living it. Wimsatt and Beardsley advocate 'language-based interpretative criticism, although they are not against taking into account the author's biography. Foucault's (1984: 102) famously pronounced the death of the author, while Hirsch (1976: 48, quoted on p. 136) insisted on reconstructing as closely as possible the author's meaning. Mitchell (2008: 144, quoted on p. 137) is inclined to view intention as a feature of the text, while Herman (2008: 237,

quoted on p. 137) claims that it is natural for a human to ascribe meanings to texts ('humans approach one another [...] as intentional systems'). The method shown in this section may be viewed as reconciling all the approaches outlined in Sotirova's overview: subliminal authorial intention seems to be recoverable by corpus means as part of its meaning.

5.7 Conclusion

The section deals with the application of Contextual Prosodic Theory to Slavic languages, particularly Russian. It proves that in these languages, as well as in English, collocation may be successfully employed to interpret the meaning of authorial text through co-selection chunking states of affairs, (fractured) prosodies and relexicalisation. Context of situation connects the textual event to similar events in the reference corpus. At the level of grammar, the most frequent quasi-propositional variables (QPVs) determine the subtext of a grammatical string and thus have a bearing on textual interpretation.

Speaking of semantic prosodies, those in Russian and Serbian which were studied in the section correspond to their English counterparts, but appear to be less monolithic. This might be explained by their diachronic development, namely, in Russian several positive collocates might have become affixed to a verb which developed a more prominent negative prosody at a later stage, similar to English 'miracles' and 'happen'.

When it comes to relexicalisation, the section shows that it is gradable and depends on the relexicalising power of the surrounding collocates. Naturally, it is more complete if the relexicalisers are used in their literal, and not in their delexical sense , and is not directly related to the number of the present relexicalisers. One more relexicaliser might make the node more delexical.

The section also undertakes to find evidence for Louw's view that all devices relexicalise. After comparing the subtext of examples taken from Yeats, Williams and Pushkin, a working corpus-attested definition of a device may be said to emerge. A ***device*** is a fractured lexical or lexico-grammatical collocation containing lexis which is relexicalised outside it. An expression which is used for the most part delexically in the reference corpus is not considered a device if its collocates or supporting context clues are

also delexical, because there is no deviation from the language norm. Since both linguistic deviation and relexicalisation may be a matter of degree, devices are likely to exist on a scale of variation of degree. This approach may point to particular instances when a device is intended, but, generally speaking, it does not aspire to raise the question of intentionality when it comes to devices. This is not to be confused with intended meaning: in the case of Williams's poem the present, chosen, lexis is not relexicalised, but the absent, most frequent QPVs are. If the context clues support the lexis chosen by the author but not the frequent QPVs, the emerging meaning is unintentional, and authorial attitude is computationally recoverable. However, without a complex study involving recoverable states of affairs, shorter grammatical strings and the author's corpus, there is a danger of an oversimplified finding instead of a valuable insight into both authorial intention and authorial attitude/meaning, distinguishable from each other.

The subtext in Pushkin's stanza, studied in this section, is not supported by the context clues – indeed, it goes against them. Therefore, the lexico-grammatical collocation under discussion may only be considered a device at the level of the lexical. If the subtext is flouted without additional context clues supporting the mismatch, the expressed meaning is inadvertent; neither can any device be intended. The subtextual recovery of authorial insincerity is now demonstrably a nascent reality.

6 Automating translation

6.1 Introduction

This chapter will illustrate how CPT can be used not only to analyse an existing translation, but also to produce a corpus-based one. The second section is a corpus-based critique of a translation of a stanza from one of the most famous poems by Alexander Pushkin. The third one shows a mechanism of producing a corpus-based translation of the same stanza, and then compares this corpus-based translation to the existing one by the renowned American translator James E. Falen. The fourth section compares the subtext of the line *Передо мной явилась ты*, extracted in the course of the previous chapter, to the subtext of Falen's existing translation of this line. This chapter is based on Louw and Milojkovic (2016: 241-272) and Milojkovic (2016b).

6.2 Analysing an existing translation

This section investigates an existing translation of one of the most famous poems by the Russian poet Alexander Pushkin. Here is the poem in the original and in its literal translation:

Если жизнь тебя обманет,	If life deceives you
Не печалься, не сердись!	Do not be sad, do not be angry!
В день уныния смирись:	In the day of despondency, submit:
День веселья, верь, настанет.	Have faith that the day of joy will come.

Now, here is Kneller's translation[22]:

If by life you were deceived,
Don't be dismal, don't be wild!
In the day of grief, be mild:
Merry days will come, believe.

[22] The translation by Mikhail Kneller can be found on http://www.kulichki.com/moshkow/LITRA/PUSHKIN/ENGLISH/kneller01.txt

It is not easy to translate Pushkin's naturally flowing trochees.[23] The third line of the stanza was chosen for investigation as it seems to carry the main weight of the poet's message. Besides, the grammatical string seems convenient for the practical purpose of investigating subtext.

6.2.1 Subtext in the existing translation

A search in the *Times* corpus yielded the following concordance for 'in the * of * be':

```
1 uld be allowed and the judgment in the sum of £160,651.16 be set aside and subst
2 a Persian minister who features in the book of Esther be hung on their own gallo
3 de to help her and, in so doing in the heat of battle, be distracted from his mi
```

Figure 6.1 The concordance of 'in the * of * be*' in the 1995 *Times* newspaper corpus

The first line's wider context is a legal case; in the second line a political speech fraught with biblical and historical allegory predicts that the enemies of the loyalists of Ulster 'would be hung on their own gallows'; the third one mentions distraction from the mission 'in the heat of battle'. The one definitely religious context points at, even if metaphorical, bloodthirstiness, the third one is set in the midst of combat; the first one falls into the same category with the following two in that a legal case may be viewed as a scene of conflict, however dispassionate the language used in the courtrooms. It may be stated with reasonable assurance that the state of affairs invoked by the search line is not one of resignation, but rather of purposeful action, out of belief in one's righteousness (in the third concordance line the male soldier is distracted from battle when helping a female comrade, which is in itself a brave and moral action). Action and moral righteousness in a conflict situation seem to be the underlying motif of the search line.

As 3 lines may not be enough, I consulted the BNC, which yielded the following concordance:

[23] The translator was certainly faced with a difficult task. Still, some linguistic choices in the stanza (e.g. 'don't be dismal') could have been improved without recourse to corpus stylistics. However, the focus here is on line 3 because the choices in it involve subtext and finer nuances of collocation.

1 n instant high profile platform in the House of Lords **be allowed** to use a TV cha
2 n instant high profile platform in the House of Lords **be allowed** to use a TV cha
3 **nor** can any grammarian **skillful** in the art of dialectic **be found to depict** this
4 **Let those things that were done in the dark of night be revealed for your justic**
5 onth before his death.   **In the name of God be it Amen, the xxthe day of**
6 es that are obviously not equal in the minds of voters **be treated** as if they wer
7 ound us, **so that** adults who are in the care of children, be they teachers, or pa
8 University can bestow and may, in the case of Nottingham, be either a Master's
9 ty. Quality of product **will**, as in the case of films, **be compared** with commercia
10 ution to a problem **can** then, as in the case of imprinting, be wired into the sys
11 **redible god. The word god must, in the minds of all, be divorced for ever from a**
12 y-unladen, chat. But it may, as in the case of Locke, be meant to be taken serio
13 **y Michael, Archangel, defend us in the day of battle. Be our safeguard against t**
14 of the warranties which would (in the absence of exclusion) **be implied** under th
15 nt, in many cases, **courts** will (in the absence of fraud) be less likely to wish
16 **for better or worse, I'm right in the middle of it. Be patient with me, for jus**
17 Art Institute of Chicago would, in the opinion of many, be wonderful for the Met
18 eserved all he got. And I'm not in the business of spy-catching. Be&rehysides, I
19 etical and **need not have arisen** in the course of or be material to an actual leg
20 **uare for the Victims of Fascism** in the capital of Zagreb **be renamed** the Square o

Figure 6.2 The concordance of 'in the * of * be*' in the BNC

In the 5 religious contexts found in the 20 lines (4, 5, 11, 13, 16), 3 are actively 'fight-picking' in tone (4, 11, 13), and line 16 is again reminiscent of action rather than inaction. Conflict is present in lines 20 (the Victims of Fascism square is being renamed), 19, and 3 (the collocates suggestive of conflict are in bold). The dispassionate language of the courtroom or attempts to solve difficulties by laying down rules is present in lines 1, 2, 6, 7, 9, 14, 15. A solution to a problem is presented in 10. Therefore, 7 lines out of 19[24] create the state of affairs of some form of conflict. Solutions are attempted in another 9 lines, including those directed by courts. Nowhere in the concordance is any form of passive response suggested. Even if the BNC concordance is not as conflict-oriented as the one from the *Times* newspaper, action underlies it to such an extent that even the tone in the remaining lines is lively and far from apathetic. Lines 8 and 17 speak of higher education (the attaining of which demands lots of action) using vocabulary like 'bestow' and

[24] Line 18 ought to be excluded as it does not actually contain the verb to 'be'.

‘wonderful’. The overall impression of the two concordances is one of lively action, occasionally argumentative or entering combat. The religious contexts are marked by righteous anger and a religious yearning.[25] In the days of grief, be mild? The case we find represented here is quite the opposite.

Two significant objections may be raised at this stage: both concordances may be coincidences, and while the first contains the verb ‘to be’ in the infinitive form followed by a past participle and a preposition in all 3 cases, the second is a syntactic mess. The verb ‘to be’ is followed by past participles, or, occasionally, adjectives, but what about lines 13 and 16? The context of both is religious and both feature a full stop before ‘be’ (‘be’ is imperative in both cases), thus containing two separate sentences. How are we to tame a concordance of such a variety of syntactic structures and are we at all justified in viewing it as a meaningful subtextual whole?

The very fact that ‘be’ in several cases begins a new syntactic entity offers a scientifically valid solution. What if we view the search line ‘in the * of * be *’ as a flexible whole consisting of two separate collocating structures: ‘in the * of *’ and ‘be *’? A claim has been made that this search line creates a state of affairs describable as activity verging on conflict. What if the order of appearance of the string's structural collocates is reserved? If the same state of affairs is preserved, we have additional proof that it is not a chance occurrence.

The search line ‘be * in the * of *’ yields 471 concordance lines in the *Times* corpus. Obviously much more frequent and thus certainly covering a broader scope of events, will it still preserve the overall impression of lively action, sometimes fight-picking or even conflict?

In the sea of 471 lines in the 1995 *Times* corpus, two random concordances, 20 lines each, were sampled. This is the first:

```
 1  that the loss was likely to be somewhere in the region of £1.6 million.   Barco
 2  that the loss was likely to be somewhere in the region of £1.6 million.   Barco
 3 "commitments" looks to me to be somewhere in the range of £5 billion to £7 billi
```

[25] It will be seen later that Pushkin’s line may evoke religious associations because of the words *унынье* (‘resignation’) and *смирись* (the imperative form of ‘resign’, ‘submit’)

4 matched companies thought to be somewhere in the range of £60 million. A more co
5 good steady state we need to be spending in the region of £900 million to £1 bil
6 y house design has tended to be rooted in the spirit of the 18th and 19th centur
7 r/ new Labour project should be launched in the spring of 1996. **Overtaken by eve**
8 annual reversionary bonus to be declared in the spring of 1996. David Prosser, L
9 on Heathrow Express, due to be opened in the summer of 1998. A further proposal
10 **judgment**, the answer was to be found in **the words of section 31**. Even though it
11 or such a declaration was to be found in **the words of section 8** which clearly co
12 that the Birlas must always be engaged in the building of a Hindu temple for th
13 s to pay off), transport can be arranged **in the back of a cross-Channel lorry**, f
14 behaviour, he **argued**, could be **justified** in the service of a prince who was mot
15 ve notice or bad faith could **be inferred in the mind of a person wholly ignorant**
16 problems because they might be needed in the event of a decision to **withdraw** th
17 been adjusted to what **might be called**, **in the fashion of a certain political pa**
18 million jackpot, which must be something in the region of a 1 in 3 billion to 4
19 ic and trust staff **could not be guaranteed in the event of a hunt** taking place,
20 tractual duty of care should be implied in the case of a domestic **tribunal**. In

Figure 6.3 The first random concordance of 'be * in the * of *' in the 1995 *Times* corpus

The collocates are less conclusive in this concordance, but wider contexts were again opened. This grammatical pattern is obviously used when discussing changes on the financial market or reporting legal judgements. There is one positive change in the religious domain (building a Hindu temple, line 12). 4 financial and one material improvement are mentioned (lines 4, 8, 18: 9). Financial losses are stated in lines 1 and 2 (the same line repeated twice). The remaining lines are either criticisms or the neutral language of litigation, on the whole amounting to 14 instances out of 20 (in fact, 19, as one line was repeated).

The second random concordance showed a similar state of affairs, supported by 16 lines out of 20, as follows:

1 ly. A further £2 billion will be available in the form of a credit that can be
2 ontracts worth £5 billion would be agreed in the course of this year. That targe
3 rally bear. **Paragraph 18 had to be read in the context of the whole document** and
4 rchase law. **The 1961 Act should be read in the light of the earlier statutes**, ha
5 x parte Barker ((1994) STC 731) be made in the course of dealing with the pendin
6 first six months of 1995 are to be published in the week of July 24, had been ex
7 e sale of the 23 airports would be delayed **in the face of growing opposition** to
8 ntinued that **section 459 had to be viewed in the context of legislation** and case

 9 out that **article 51 could only be invoked in the case of actual armed attack**.
10 y 6**-7, 6-7, 6-3, 6-4, 6-3**, will be etched in the memory of all who witnessed it.
11 laboratory. Such a sensor could be buried in the skin of the arm and attached to
12 s enable up to £3,000 a year to be invested in the shares of one firm. When take
13 e tickets, a call for **a line to be drawn in the shag-pile of the VIP departure** l
14 ffered by General Abacha had to **be seen in the context of pressure** from a group
15 es to players about **bribes, may be "jocular"** in the eyes of one former Pakistan
16 r **human rights abuses**. This may be mainly **in the hope of regaining credibility** a
17 ut even that achievement has to be seen **in the context of a second hefty defeat**
18 air. Today every **actor seems to be moving in the direction of direction**. Robert
19 "although", he added, "it could be **obstinate in the face of threats**." B.B.C.
20 xpensive **legal aid lawyers**, may be high in the list of considerations. In this

Figure 6.4 The second random concordance of ‘be * in the * of *’ in the 1995 *Times* corpus

Here, 3 positive pieces of news in the domain of finance (1, 2, 12) and one medical breakthrough (11) leave 4 lines of legal language (3, 4, 5, 8), 3 military contexts (9, 14, 16), 3 sports ones, one of them involving attempts at the bribery of players (10, 15, 17), 4 critical and/or ironic ones (13, 16, 18, 20) and one obviously defiant (19).

The following conclusions obtain:

- grammatical strings create states of affairs, consistent in approximately two thirds of the lines
- these states of affairs are not counteracted by the remaining lines (e.g. the reports on financial losses or praise of universities do not oppose the overall impression of righteous action in that they do not suggest resignation, the state of affairs looked for in the original)
- longer grammatical strings creating subtextual states of affairs may be broken into grammatical ‘collocates’; their reversed order creates the same or very similar states of affairs

It may very well be the case that every grammatical pattern in English carries with it a certain kind of semantic aura, more or less definite, and it is certainly worth investigating. This also seems to be the case with Russian (Section 5), and should be the case with languages where semantic prosodies have been found, like Serbian, German and Italian.

All the same, it would be unrealistic to issue a general recommendation to writers and translators to consult reference corpora every time they write a line of text. What is being proved in this dissertation and has been elsewhere (e.g. Louw and Milojkovic 2014) is the fact that certain *key* lines in original poetry may and do contain subtextual meaning, similar in nature to semantic prosody. If we are dealing with translation, this key line may be intuitively rendered into the target language with the help of a grammatical string containing *opposite* or in some other way unhelpful subtext. Moreover, it sometimes happens that the author may unconsciously (or for lack of inspiration) use a grammatical pattern, the subtext of which (traced in the reference corpus) differs from the author's intended meaning. These subtextual clashes are dealt with elsewhere in this dissertation (see sections 5 and 9). Such clashes suggest that corpus-based instrumentation might be applied to the translation process. When it comes to subtext, much is to be gained by looking at the semantic relations between the lexical variables in the corpus and those of the author's/translator's choice.

Although the terms 'subtext' and 'subtextual clashes' were used in the previous passage to denote states of affairs created by a certain grammatical pattern, our discussion of 'in the * of * be' has so far dwelt on nothing more than semantic prosodies, 'auras' of meaning surrounding a word or phrase and recoverable through large reference corpora. Ideally, subtext is in fact recoverable through content words that are wildcarded in the search of a grammatical string, and not through the general aura of meaning surrounding it. The quasi-propositional variables of the search line help in determining the meaning of the line. In our case, this at least means that if the translator is dealing with a line containing advice to be resigned, the subtext ought not to contradict resignation unless the poet intended it in the original. Therefore, subtext, or wildcarded quasi-propositional variables in the place of 'day' in 'in the day of grief be mild' must not prospect conflict or action aimed at solution, but ought to prospect inactivity and resignation.

As too few lines containing 'in the * of * be *' are yielded by corpora, we must keep our 'grammatical collocates' hypothesis' and reverse the word order of the search line, making it 'be * in the * of *'. The most frequent QPVs of this search line are context (43

occurrences), light (36), case (18), middle (17), form (16), event (12), house (9).[26] To be scientific, we will need to study all the variables to (dis)prove our assumption that this grammatical string is a call for action in what could even be a conflict situation.

Let us begin with the less frequent QPVs of this grammatical string. The least frequent lexical variable, 'house', refers to none other than either of the two Houses of Parliament. This in itself sets the scene for contexts of debate:

1 wners and could probably only be **challenged** in the **House of Lords**. "On the bas
2 d **case on eavesdropping is to be heard in the House of Lords**. Sultan Khan, ser
3 mber, when the programme will be recorded in the House of Commons for the first
4 rting **naval guns** to Iran will be **raised** in the **House of Commons** tomorrow, when L
5 earer, a simple answer was to be found in the House of Lords' decision in The Fa
6 num. The Pensions Bill will be **debated** in the **House of Commons** after Easter. M
7 t kind. Benson continued to be active in the House of Lords, however, speaking
8 the Haymes Garth **affair** will be placed in the House of Commons library this aft
9 ill effect this merger, is to be **debated** in the **House of Commons** on April 19, it

Figure 6.5 The concordance of 'be * in the house of' from the 1995 *Times* newspaper corpus

The collocates 'debate' (appearing twice), 'challenged' and 'raised' all point to conflict-based situations, while the seemingly neutral collocate 'heard' collocates with 'eavesdropping'. Our second least frequent QPV is 'event'. The collocates in bold, directly or indirectly pointing to emergencies and safety risks, are found in every line:

1 ons of water. Also, they will be self-righting in the event of **capsize**. Compet
2 tures of the scene. They will be invaluable in the event of an **insurance claim**.
3 said Young Hustler would only be withdrawn in the event of an **overnight deluge**.
4 lic and trust staff **could not be guaranteed** in the event of a hunt taking place,
5 ed that **all prisoners were to be executed in the event of an Allied invasion**. Th
6 e **problems** because they might be needed in the event of a decision to **withdraw t**
7 request for a **hearing** date to be fixed in the event of an extension being **refuse**
8 ould be substantial wealth to be shared in the event of Mr Ranno's **death**. This i
9 that their **Marines** would only be used in the event of a **crisis where UN troops w**
10 ts and pets that will have to be **evacuated in the event of an emergency**. The que
11 t on short-notice stand-by to be sent in the event of **a breakdown in security**.

[26] The QPVs come from the same *Times* 1995 concordance, consisting of 471 lines.

12 he Rapid Reaction Force could be used in the event of a Serb **attack** on the vital

Figure 6.6 The concordance of 'be * in the event of' from the 1995 *Times* newspaper corpus[27]

While the previous concordance pointed to debate, this one is decidedly permeated with danger to one's health and often even life (line 2, the least lethal-sounding, is taken from an article entitled: 'What should I do if I come across an accident on the road?'). Two of these situations do not call for quick solutions (definitely not line 8, while in line 3 'deluge' is used ironically), but the remaining 10 contexts do.

Will 'in the form of' offer more placid states of affairs? As it turns out, there is no danger to one's life in the following concordance, but neither is there any particular reason to remain passive. Rather, 'in the form of' is mostly about any form of (mostly financial) problem-solving:

1 en where that will had yet to be expressed in the form of **law**. Whether by accide
2 ed gravitational lensing, can be detected in the form of split images of quasars
3 . A further £2 billion will be available in the form of a **credit** that can be u
4 of Los Angeles **cannot easily be captured** in the form of the traditional novel.
5 ple flee. If so, there **should be evidence**, in the form of glassy **volcanic fragme**
6 , comes through and will also be interested in the form of Ian Arnold, the forme
7 tober 30. The **new shares** will be placed in the form of American Depositary Share
8 exists for **easing policy** will be taken in the form of **tax cuts**. It is, therefore
9 as cash) and the **balance must** be taken in the form of an annual income, which is
10 weaks. But the British will be there in the form of the revitalised **Triumph**. O
11 uarantee for all (which would be taken in the form of **tax relief** for those in wo
12 . **Financial support** will also be made in the form of a **grant** to fund constructio
13 April 1987) the **benefits must** be taken in the form of income, not as a lump sum.
14 t of the **wage rise** would soon be felt in the form of **rationalisation** and further
15 y way of **relief** would usually be **granted** in the form of an interim injunction.
16 er than the lump sum, have to be taken in the form of an annuity which is taxed

Figure 6.7 The concordance of 'be * in the form of' from the 1995 *Times* corpus

[27] Line 6 says: '…in the event of a decision to withdraw the 25 000 UN troops.'

The third least frequent second QPV of ‘be * in the * of’, form, has not produced as many debate-fraught or danger-fraught context. On the contrary, exactly half of the lines speak of financial offerings, e.g. in the form of a credit (line 3), tax cuts (line 8) or a grant (line 12). These are financial solutions, as well as ‘new shares’ (line 7), an annual income (line 9), tax relief (line 11), benefits (line 13), and an annuity (line 16). Line 14 is only deceptively positive (the wage raise will lead to redundancies), as its wider context is a strike being solved through negotiations. The other negative lines are line 1 (about a debate surrounding the EU) and perhaps line 2 (the possibility of a volcanic eruption in the past that had made a population flee). Contexts 2 and 10 are positive but not calling for any action in particular. All in all, resignation does not come into the states of affairs created by this concordance.

The next QPV to be examined is ‘middle’. 7 lines out of the given 17 refer to middle as a point in time, but are their contexts suggestive of passivity?

1 elvyn Bragg **does his best to be sympathetic. In the middle of his questions** he u
2 ed that **Hitler's body should be destroyed. In the middle of the night** of April 4
3 t frostbite, just as he will be there in the middle of spring to see that nest-e
4 a proposal, it was likely to be made in the middle of October. The protracted
5 ut if they are henceforth to be held in the middle of winter, you may rest assur
6 , "I find it **very strange to be stuck in the middle of something where I am tryi**
7 <Group> HAILING a cab is **to be banned in the middle of Manhattan**. From next sum
8 was not uncommon for them **to be telephoned in the middle of the night and summon**
9 a three-hour quiet period to be observed in the middle of the day. The high le
10 recorded Hogmanay special to be broadcast in the middle of a Goldie Hawn film. O
11 nitoring systems, which will be completed in the middle of the year," Earl Howe
12 The contract is expected to be signed in the middle of next month and engineeri
13 of personnel functions would be completed in the middle of the year, including p
14 lished that **the rig will not be sunk in the middle of the Atlantic**, it is essent
15 ent to bed at last, I was to be **awakened in the middle of the night by the leaki**
16 ternational troops **who would be caught in the middle of a bloody battle** for Sara
17 rnment **decided she could not be dismissed in the middle of her court challenge** a

Figure 6.8 The concordance of ‘be * in the middle of’ from the 1995 *Times* newspaper corpus

The expressions with ‘middle’ meaning ‘a point in time’ (lines 3,4,5,6,11,12,13) seem to depict situations in which an active solution is being sought, even if the situations themselves are not as urgent as, say, in the remaining lines of this concordance. Here the lines containing ‘in the middle of the night’ (2, 8 and 15) are not included, as the sense of urgency overrides the purely temporal meaning of ‘night’ (in lines 2 and 8 it is the sense of secrecy as well as urgency, but also of action). As for the rest, line 1 is certainly about melancholy (that of Martin Amis), but it is the only such line. Line 6 (‘stuck in the middle of something’) is argumentative. So is the wider context of lines 7, 10 and 14. Context 17 depicts a conflict, while context 16 warns of life danger to international troops. All in all, only line 1 of the 17-line concordance is suggestive of passivity and resignation.

So far, the less frequent QPVs (‘house’, ‘event’, ‘form’ and ‘middle’) seem to conform to the overall semantic aura of the grammatical string. The ‘Houses’ referred to are the one of Lords and the one of Commons, where debates are expected to occur, as they do in the concordance. The ‘event’ concordance supports our intuitive assumption that the event in the grammatical string ‘be * in the event of’ is that of emergency, demanding an urgent solution. Even ‘in the form of’ is mostly about (financial) problem-solving, while ‘middle’ calls up 17 contexts, only one of which describing lack of action.

Let us look at the concordances of ‘case’, ‘light’ and ‘context’, which are here given in full. When it comes to ‘case’, whether it is more or less evident to the reader who cannot open the wider contexts, the majority of the lines refer to the courtroom:

1 **uction and distribution would be controlled. In the case of tablets** such as Ecst
2 Capital growth is unlikely to be significant in the case of bond Peps, indeed so
3 akes it **unlikely that it will be successful** in the case of other products. This
4 f what had happened, it would be advisable in the case of juveniles as a matter
5 t division Though looks can be deceptive, in the case of Pat Lam, the engaging
6 **Reporting restrictions** should be honoured in the case of Mrs West, to prevent th
7 eter. **Radiocarbon dating** will be **supplemented** in the case of the oak samples by
8 tion. The requisite **intent** to be proved in the case of **murder** was an intention t
9 ut that **article** 51 could only be invoked in the case of **actual armed attack**. He
10 a **circumstance** that could not be present in the case of the **hypothetical man**.
11 e and that should be taken to be present in the case of the **hypothetical man** had
12 a circumstance that could not be present in the case of the **hypothetical man**.
13 . What then was the test to be applied in the case of a **public interest petiti**

14 re. The problem was said to be that in the case of a **concealment** taking place
15 sonable cause. That would not be possible in the case of genuine **complaints**. I
16 tory-made parts was deemed to be **dicky**. In the case of some **"permanent" heavy** co
17 ntractual duty of care should be implied in the case of a domestic **tribunal**. I
18 of this possible **trend** is to be seen in the case of **members of the House of Com**

Figure 6.9 The concordance of 'be * in the case of' from the 1995 *Times* corpus

Namely, the states of affairs of the courtroom are present in lines 4, 6, 8, 10, 11, 12, 13, 14, 15, and 17 (10 lines out of 18). Additionally, lines 1, 2 and 3 are argumentative, and line 9 discusses military issues. There are 2 positive lines (5 and 7) and 2 contemplations (16 and 18, the latter not without an argumentative tinge). Nothing at all in this concordance suggests passivity.

The next string to be discussed is 'be * in the light of'. Unsurprisingly, the variables in the first lexical slot for the most part suggest steps which are to do with drawing conclusions before taking action:

1 skinned vessels might have to be **abandoned in the light of the Derbyshire invest**
2 ad, in the present **case**, also be **considered in the light of article 10**: see Youn
3 accident? The **issue** had to be **approached in the light of the trial judge's fi**
4 1/26). The **regulations** had to be **construed in the light of the wording and purpo**
5 MI5's £150 million **budget** to be **increased in the light of its expanded role**.
6 young women, are all seen to be explicable **in the light of predictions derived**
7 **project** clearly would have to be **considered in the light of so many other demand**
8 r national **legislation** should be **interpreted in the light of later unimplemented**
9 grow up. But all **rules** should be **simplified in the light of experience**; many cou
10 a segment of the **trial** would be **allocated in the light of the anticipated needs**
11 ment of a final **dividend** will be **"reviewed in the light of second half and full**
12 xercise of that **right**, had to be **determined in the light of the activities which**
13 respect. Their **resentment** can be **understood in the light of the level of casualt**
14 ment and how the deer were to be **conserved in the light of it. His Lordship** wo
15 **exercise of that right** had to be **determined in the light of the activities which**
16 ry nor desirable", has now to be **seen in the light of Hurd's reluctance to vacat**
17 hase law. The 1961 Act should be **read in the light of the earlier statutes**, havi
18 of course **news judgments** must be **made in the light of other stories** on any parti
19 **itability of candidates** would be **assessed in the light of published criteria**, Lo
20 ment of a final **dividend** will be **reviewed in the light of final results** for 1995
21 nsultation and will certainly be **amended in the light of what we are told,"** a sp
22 ommended that the **plan** should be **shelved in the light of the recent killings** in

23 le European **currency** can only be **made in the light of the, presently unforeseeab**
24 rn oil frenzy in Alaska could be **seen in the light of the gold mania 100 years a**
25 stem,but hoped the **case** could be **reviewed in the light of new evidence.** Mr Rom
26 ision and the **decision** had to be **re-taken in the light of his Lordship's judgmen**
27 This matter would inevitably be **raised, in the light of John Smith's death** last
28 ility of nervous shock was to be **judged in the light of what would be suffered** b
29 a single **currency.** This would be **risky in the light of strong anti-EU sentiment**
30 rint, that performance should be **measured in the light of cyclical conditions.** T
31 . Yesterday's action should be **seen in the light of the London conference's ul**
32 he **decision to proceed** had to be seen in the light of the prevailing climate, ma
33 ps our **integrity** now needs to be **assessed in the light of what we are making of**
34 favoured option but this may be **reviewed in the light of the Brent Spar debacle**
35 t the **bargaining process** must be **altered in the light of global realities.** "If w
36 he **decision to proceed** had to be **seen in the light of the prevailing climate,** ma

Figure 6.10 The concordance of 'be * in the light of' from the 1995 *Times* corpus

The variables in the first lexical slot range from those denoting judgement rather than action (e.g. 'seen' (5 occurrences), 'reviewed' (2 instances), 'considered', 'assessed', 'understood') to those denoting action (' abandoned', 'altered', 'conserved', measured', shelved'), with a variety of variables in between these two extreme points on the scale (e.g. 're-taken' collocating with 'decision' or 'raised' collocating with 'matter'). Although one might expect the expression 'in the light of' to give its concordance more contemplative a character, still, passivity or resignation does not underlie any of its lines.

Finally, we have come to 'be * in the context of'. This is the concordance from the *Times* 1995 corpus, given in full. The absolute majority of the lines contain a modal or semi-modal followed by the present passive infinitive, like 'be *ed', which is in many instances indicative of action:

1 **netary policy is supposed to be set** in the context of achieving inflation at 2.5
2 t any stock provisions would be "immaterial" in the context of the value of the
3 upportive" spirit and **should be seen** in the context of a much broader campaign t
4 r's **call that monetary union be viewed** in the context of making Europe more comp
5 red by General Abacha **had to be seen** in the context of **pressure** from a group of
6 epidemiological evidence **to be interpreted** in the context of her personal circu
7 the north of the Sahara **will be considered** in the context of the whole continent
8 . All these **failures have to be set** in the context of her **reputation for loyalty**
9 erry's comments, they **should be seen** in the context of what has been an extraord

10 is year's European Tour **will be played** in the context of the **struggle** between Mo
11 recedes, **the event tends to be judged** in the context of the **war**'s more orthodox
12 ly bear. Paragraph 18 **had to be read** in the context of the whole document and of
13 bassadorial role abroad **must be addressed** in the context of her future position
14 eet is ecumenical." This may be true in the context of **a pilgrimage** through Tami
15 that the winning entry **could be seen** in the context of a wider selection of the
16 lth of **the Falklands need to be read** in the context of Argentina's **fractious** pol
17 **only England player** who **can be mentioned** in the context of **genius** without court
18 or Christians **the issue will be approached** in the context of their relationship
19 e arts. All of this **should be seen** in the context of popular culture's extraor
20 n its own right. It **can only be effective** in the context of another medical or p
21 Those credits, however, **must be seen** in the context of a side where every other
22 t authorities and **should not be followed** in the context of the modern rules of c
23 hat **business strategies will be considered** in the context of its brand portfolio
24 ." Current results **have to be viewed** in the context of the level she is compet
25 itish queen" **must**, moreover, **be seen** in the context of her **pique** at her own comp
26 om **the axe has fallen should be seen** in the context of a group with a UK head-of
27 even that **achievement has to be seen** in the context of **a second hefty defeat** and
28 . This instance may indeed be **useful** in the context of the **crisis facing** the C
29 sk. Those questions **had to be asked** in the context of **who was responsible for**
30 he **must explain how this can be done** in the context of local financial managemen
31 g that the invitation **should be seen** in the context of conversation throughout t
32 blican objectives **could even be discussed** in the context of all-party talks. T
33 language **only if they could be justified** in the context of the work. "This wi
34 nued that section 459 **had to be viewed** in the context of **legislation** and case la
35 s or **the head of loss gad to be considered** in the context of the cause of action
36 uences. **Tax proposals must be seen** in the context of the tax system as a whole
37 sed "questions which **have to be considered** in the context of this whole constitu
38 ckets and making runs **has to be put** in the context of a genuinely friendly match
39 n innings: all these **have to be put** in the context of the man's scoring rate, 5.
40 1984, a decision that **should be viewed** in the context of the Government removing
41 The whole **debate** now **has to be set** in the context of the internecine war which
42 gued that any **errors need to be put** in the context of their broader record. Pr
43 7 million in November **has to be put** in the context of total outstanding consumer

Figure 6.11 The concordance of 'be * in the context of' from the 1995 *Times* newspaper corpus

Similarly to 'be * in the light of', this concordance contains QPVs that involve contemplation rather than action (e.g. there are 10 instances of 'seen', 4 instances of 'considered', 4 instances of 'viewed', 4 instances of 'put' and 3 instances of 'set'). These

QPVs predominate (29 in 43 lines), followed by 6 QPVs denoting a point between contemplation and action (‘interpreted’, ‘judged’, ‘mentioned’, ‘asked’, ‘justified’, ‘discussed’). Action is expressed by 7 QPVs (‘played’, addressed’, ‘approached’, effective’, ‘followed’, ‘useful’, ‘done’). Despite the fact that the QPVs in the first lexical slot do not denote action in their majority, contemplation (as in ‘seen’, ‘considered’) might be viewed as activity of the mind. Besides, the modals and semi-modals denoting action predominate (13 forms of ‘have to’, 7 occurrences of ‘should’, 4 of ‘will’, 3 of ‘must’ and one of ‘need’: overall 28 occurrences out of 43). It is evident that the concordance as a whole does not suggest resignation.

Now that we have observed the role of the most frequent QPVs in the studied grammatical string, let us imagine them fill the second empty slot in ‘be * in the * of’, rather like in a slowed down cartoon:

be * in the **context** of
be * in the **light** of
be * in the **case** of
be * in the **middle** of
be * in the **form** of
be * in the **event** of
be * in the **House** of

Rather than attempting to discover one underlying feature the QPVs share, we should observe how each two QPVs share a meaning. Let us imagine the first two QPVs (or the next two, or any two) as if in a dialogue together.‘In the context of’ and ‘in the light of’ are virtually identical in terms of meaning. ‘In the light of’ and ‘in the case of’ suggest (some aspects of) a situation. ‘In the middle of’ and ‘in the form of’ suggest a solution in a specific form to be administered in the middle of a certain situation… And so on.

‘Day’ does not appear on the list of the QPVs of the studied grammatical string. Now let us imagine the string ‘be * in the **day** of’. What is it that any pair of the QPVs does not share with ‘day’? ‘Day’ is static. A point in time. It does not point to a situation or its

specific features. As it indicates a point in time, it corresponds only to the QPV 'middle' in its temporal sense.

Therefore, 'in the day of grief be mild' is a grammatical string whose overall semantic prosody and subtext point in the direction of action in a specific situation (sometimes conflict), whereas the meaning this grammatical string is supposed to convey is one of resignation. The translated line also includes a content word ('day') in place of QPVs which suggest motion. Let us check a similar grammatical string where 'day' *does* appear as one of the collocates – 'be * **on** the * of'. The most frequent collocates of the searched grammatical string in the second lexical slot are: 'basis' (31), fingers (7), ground (7), back (7), eve (6), day (5).This is a section of the 250-line concordance, chosen because of the presence of the QPV 'day':

231 tal role which certainly **cannot be discharged on the basis of** a meeting a year.
232 l Commons committee **rather than be discussed on the floor of the Commons**. They c
233 wherever it comes Gunn **seems to be definitely on the side of** the glossy magazine
234 loyalty to the monarchy and **to be circumspect on the prospects of a divorce**. Th
235 ublic. Open government **need not be sacrificed on the altar of efficiency**. <
236 ding exchange of contract, **must be completed on the day of the sale**. The selle
237 was said to be **too sensitive to be disclosed on the basis of** public interest imm
238 most of the characters **seem to be constantly on the edge of financial collapse;**
239 llet matinees **can fairly easily be scheduled on the day of** an evening performanc
240 on 37, their **decision could not be invalidated on the ground of** an improper moti
241 itish Gas last night **refused to be drawn on the timing of any rise**, adding that
242 at 20p per share, but **would not be drawn on the timing of any future increase**.
243 hill estate that £500,000 **is to be spent on the upkeep of the grave**, at Bladon i
244 nclude that restrictions **should be placed on the use of typewriters** because of t
245 ew limit of £100,000 is also **to be placed on the value of houses and mortgages** w
246 am had died." That entry **should be pinned on the wall of** every dressing-room in
247 tingham Panthers. The ties **will be played on the weekend of** October 14 and 15.
248 ictory over Japan, VJ Day, **will be marked on the weekend of** August 19 and 20 by
249 the same spot. The matches **will be played on the weekend of Saturday**, March 11.
250 ndate **the death penalty and can be enforced on the word of a single witness** is m

Figure 6.12 A section of the concordance of 'be * on the * of' from the 1995 *Times* newspaper corpus

This part of the concordance consists of 20 lines that do not at all suggest inaction. These and expanded contexts show that 3 lines refer to sports (lines 246, 247, 249), 4 lines refer to legal cases (236, 237, 240, 245), 4 contexts are critical (233, 238, 139, 250) and 5 lines are argumentative (231, 232, 235, 243, 244). Context 248 is military. On the whole, ‘conflict’ situations (17 out of 20 lines) are not fewer than in the case of ‘be * in the * of’ (in the 2 random concordances of this line there were 14 and 16 conflict-based lines, respectively).

Here are the lines with the most frequent QPV, ‘basis’:

1 e **plaintiff's damages were to be quantified** on the basis of the loss resulting f
2 ther the **limited funds should be dissipated** on the basis of first come first ser
3 that **so big a levy could not be justified**. On the basis of projections in their
4 s said to **be too sensitive to be disclosed** on the basis of public interest immun
5 l **role which certainly cannot be discharged** on the basis of a meeting a year. Bu
6 **ash's own Euro-phobia were to be questioned** on the basis of the speeches he made
7 ame work at piece rates **shall be calculated** on the basis of the same unit of mea
8 redominantly of women, **was to be calculated** on the basis of the same unit of mea
9 promises that services **would be franchised** on the basis of present timetables w
10 ies, **but would expect this to be done on the basis of need**." If that is not a
11 e asks. **"Programmes have to be made** on the basis of the cultural significance
12 e next 18 months **will largely be decided** on the basis of figures gathered over t
13 the **tournament survival would be decided** on the basis of which country had incur
14 The **participating teams would be decided** on the basis of arena size and quality
15 tner. Occupation orders **would be granted** on the basis of who had rights to occup
16 some of its **undergraduates to be accepted** on the basis of an interview and schoo
17 tions to run a multiplex **will be assessed** on the basis of the speed and geograph
18 r 29. Free warrants **will also be issued** on the basis of approximately one for ev
19 of 12.675 million **shares will be made** on the basis of six new shares for every f
20 the means by which **care could be provided** on the basis of individual need rather
21 ion. Damages, he said, **should be assessed** on the basis of a sale at that figure
22 mands and that **to do so would be "unsound** on the basis of any principles of corp
23 exact opposite of **what should be occurring** on the basis of any fundamental measu
24 **distribution thereof that may be possible** on the basis of common ownership of th
25 those applying for help **will be chosen** on the basis of application forms alone,
26 ir condition. Payments **will be assessed** on the basis of the suffering and loss
27 ent that the franchises **would be awarded** on the basis of the present timetable.
28 destroyed. The **issue was to be decided** on the basis of the wording of the sect
29 s but enables **a comparison to be made** on the basis of the services they offer.
30 ote) is **demanded**, then it **can be arranged** on the basis of the shareholders physi
31 Gettysburg "'THE **draft will be conducted** on the basis of a lottery,' Noonan ex
32 **heritance of an estate has to be decided** on the basis of exact time of death.

33 un at Epsom or Chantilly **will be made** on the basis of all the facts and it will
34 icated that **charges would not be brought on the basis of these tests unless** ther
35 when **it was proposed that it be done** on the basis of the established views of t
36 h the border closure **can only be judged on the basis of hard evidence**. Preside
37 **unlawful-killing verdict will be returned** on the basis of evidence which might b
38 ntments at every level **should be made on the basis of merit**. But how can recru

Figure 6.13 The concordance of 'be * on the basis of' from the 1995 *Times* newspaper corpus

The whole concordance shows that in fact important decisions are made 'on the basis of…' The absolute majority of the QPVs in the first lexical slot denote action, and not mere contemplation. 13 lines (approximately one third) out of 38 are in some degree conflict-based (1-6, 10, 13, 14, 30, 32, 34, 37): this conclusion was reached on the basis of their collocates in the 9-word window. The *Times* concordance containing 'ground' instead of 'basis' is conflict-based in its entirety:

1 enforcement notice could not be challenged on the ground of irrelevant matters,
2 37, their decision could not be invalidated on the ground of an improper motive:
3 Goatley said the case would be brought on the ground of abuse of process, claim
4 cation for the indictment to be stayed on the ground of abuse of process because
5 ision that the remedy should be refused on the ground of equity. In that regar
6 the board's decision had to be reduced on the ground of illegality as explained
7 e Commission,would therefore be annulled on the ground of infringement of essent

Figure 6.14 The concordance of 'be * on the ground of' from the 1995 *Times* newspaper corpus

If 'on the basis of' and 'on the ground of' are similar in meaning, and in fact form a cline, a gradation, with 'ground' used in, it seems, exclusively conflict-based situations, what is the collocate 'fingers' doing in the subtext of our search line? It turns out that all the contexts except for line 4 (which is actually praise) are highly argumentative:

1 s in the subcontinent can now be counted on the fingers of one hand; the rest ha
2 y certified audit firms could be counted on the fingers of one hand. But, of c
3 S clinics dealing with it can be counted on the fingers of two hands, although i
4 ber of passes he wasted could be counted on the fingers of one mitten. As Wilk
5 prisoners' "rights" still to be counted on the fingers of one hand after a hund

6 e greater; but such stars can be counted on the fingers of one badly mutilated h
7 st 30 years a list that could be counted on the fingers of two hands that probab

Figure 6.15 The concordance of 'be * on the fingers of' in the 1995 *Times* newspaper corpus

What is the difference between 'on the basis of' and 'in the context of'? Why have I just used 'on the basis of the collocates' and not 'in the context of the collocates'? Especially since 'context' is what collocates are about?

'Basis' is more definite, more finite. 'Basis' is the last step before reaching a decision. This is precisely what the 'basis' concordance illustrates. 'Context' contains clues that will point into the direction of a decision. 'Basis' is static, 'context' is dynamic. Both are used in situations, sometimes conflict-based, where a decision is called for.

The other frequent QPVs are 'eve' and 'day'. In the following concordance of 'eve', there is a danger to the well-being and lives of individuals in 3 lines out of 6 (lines 2 and 4, the wider context of line 5), and in line 1 to the health of a horse:

1 or the hurdlers' crown **only to be withdrawn on the eve of the race after hurting**
2 the Dutch resistance, **only to be arrested on the eve of his departure for Engla**
3 s of a parallel IoM survey, **to be released on the eve of** next week's Conservativ
4 **survive the entire war only to be murdered on the eve of the liberation**. The n
5 **amantly opposed**. He would thus **be left**, on the eve of a presidential election ye
6 year, the **Dubai World Cup will be staged on the eve of** the Grand National meetin

Figure 6.16. The concordance of 'be * on the eve of' in the *Times* 1995 newspaper corpus

What interests us most is the variable 'day', because it appears in the target line. The contexts of day are far from static:

1 ing exchange of contract, **must be completed on the day of the sale**. The seller
2 let matinees **can fairly easily be scheduled on the day of** an evening performance
3 ng the digital revolution **will be held on the day of** publication, Thursday April
4 much **your pension fund will be worth on the day of your retirement**. Imagine y
5 ime. But **will such a request be forthcoming**? On the day of the release of the

Figure 6.17 The concordance of 'be * on the day of' from the 1995 *Times* newspaper corpus

‘Day’ is not on the list of the lexical variables in ‘be * in the * of’, but in ‘be * on the * of’. Apart from that fact, there is no sufficient basis to claim that there is a very substantial difference in the semantic auras of the two strings. The two random concordances of the former proved as conflict-based as a random concordance of the latter containing, among others, the variable ‘day’, and both promote action rather than passivity. The search of the frequent variables showed that they can be more or less suffused with action (‘on the basis of’ is more so than ‘in the context of’) or with a sense of emergency or danger (‘in the event of’, ‘on the eve of’), or with that of debate (‘in the case of’, ‘on the ground of’), depending on which variable we are discussing.

A native speaker’s intuition, however, will point at examples when ‘in the day of’ is used. The claim that it is not irrelevant whether to use ‘on’ or ‘in’ in front of ‘the day’ is confirmed if we look at a biblical context:

> Seek ye the LORD, all ye meek of the earth, which have wrought His judgment; seek righteousness, seek meekness: it may be ye shall be hid in the day of the LORD’S anger (Zephaniah 2:3).

It is obvious from the above example that the context of God’s anger on Judgment Day is too dramatic for the static preposition ‘on’. In Pushkin’s original line, the impression created is of static and not dynamic nature, due to the words *уныние* (‘despondency’) and *смирись* (‘resign’).

In order to pursue this point, let us look for the string ‘be * in the day of’ in other corpora. The BNC does not contain it, but the Google Books – UK corpus does. The 5-word string ‘be * in the day of’ was entered. The QPVs collocating with this string followed by ‘of’ in this corpus were ‘saved’, ‘willing’, ‘hid’, ‘blameless’, ‘done’, ‘given’, ‘strong’, and ‘firm’. Apart from the last two variables, the rest appear exclusively in the context of Judgement Day and of the coming of Jesus Christ. The last two appear in different contexts, quoted below:

May I **be strong in the day of battle;** may I not forget that I am a child of Deity — a humble instrument created for work! (**Charles Fransis Hall, *Life with the Esquimaux*, 1864)**

But friendship […] therefore requires […] that they should not only **be firm in the day of distress,** but gay 'in the hour of jollity; not only useful in exigencies, but pleasing in familiar life; their presence should give cheerfulness as well as courage, and **dispel alike the gloom of fear and of melancholy.** *(The Rambler*, 1750, no.63)

The Google Books – UK corpus search shows that 'be * in the day of' mostly appears collocating with Judgement Day. In the rare examples where it does not, it collocates with 'battle' and 'distress'. Although the string 'be * on the * of' seems as conflict-based as 'be * on the day of' in the *Times* reference corpus, the lexico-grammatical collocation 'be * in the day of' appears less suitable as it creates dynamic rather than static contexts.

On the basis of what we have so far discovered, semantic prosodies of grammatical strings point to certain prevailing meanings or usages. However, and in particular in the case of structures containing 'of', whose usages are bound to have a wide scope, the QPVs are the clue that helps to distinguish between the differences of usage. 'Day', except in biblical contexts and very rare examples, belongs to 'be * on the * of', and not to 'be * in the * of, as the reference corpus tells us. 'In' might denote process, and 'on' more finite steps, according to the previously discussed concordances. Nevertheless, as the second lexical variable of 'be * on the * of', 'day' denotes action and not inaction.

6.2.2 States of affairs in the existing translation

Let us now consider the states of affairs achieved in the translated text by co-selection. The adjective 'mild' is obviously chosen largely to keep the rhythm and rhyme. 'Mild' appears to fit the meaning of 'submit' or 'resign yourself' in that it is the opposite of 'aggressive'. 'Be mild' may be understood as 'do not take violent action'. However, the choice 'be mild' does sound a little unnatural. The beauty of Pushkin's lines lies in the natural flow of his trochees that ring true to every native speaker of Russian, and it is vital

that his verse retain its ease in translation. This is the concordance of 'mild' from the *Times* reference corpus:

```
 1 D: "Controversial" is altogether too mild a word to describe the Bernard Haitin
 2  of a few years back."   Major is too mild a man to be condemned as a "madman in
 3 iety."   The snag, if that is not too mild a word, is that many computer clocks
 4  to the Falkland Islands would be too mild a punishment for her support of John
 5 ce of some other cheese was also very mild. A 'warm salad of fish and chips' was
 6  and clients."   And the result? Have mild-mannered accountants metamorphosed in
 7 . Believe it or not, car washes use a mild acid to remove brake dust from wheels
 8 ve chosen publicly to attack than the mild-mannered Adrian Coles, Director-Gener
 9 nitely seeing more cases of unusually mild adult chickenpox, some of which have
10 pier is a ship without seasickness, a mild adventure. When I lived briefly in Br
11 re said last night that today will be mild after heavy overnight rain, but north
12  could range from total disability to mild after-effects followed by complete re
13 fidence women drivers displaying even mild aggression at the wheel are a rarity.
14 uires a choice of identites. There is mild-mannered Allan Shiach, heir to Macall
```

Figure 6.18 The concordance of ‘mild’ in the 1995 *Times* newspaper corpus

True, ‘mild’ does not often describe human beings, as the rest of the 676-line concordance will show, but occasionally it does. Let us see the full concordance of ‘mild’ co-selected with ‘man’:

```
 1  of a few years back."   Major is too mild a man to be condemned as a "madman in
 2  morning I saw a man embarking upon a mild altercation with a traffic cop: in a
 3 . He is a man of rich contradictions: mild and mumbling in manner but white-hot
 4  from behind the curtain. No longer a mild-mannered, diffident young man, but a
 5 touring his laboratories with another mild man who does not. Theirs is a very Br
 6 nel4's polemical series, brought us a mild man who cuts up rats touring his labo
 7 ther sporting stereotype, that of the mild man who becomes a raging bull on the
 8 rt in coming up with Lord Nolan. This mild man and his anodyne committee would c
 9  actor Tom McCamus is riveting as the mild man led off the rails by his lust for
10  her husband were unfounded. He was a mild-mannered man devoted to his family."
11 erly devastated by this case. He is a mild-natured man whose life has been destr
12 g Goldstein needed: another battle. A mild-mannered man forcing through a massiv
13 t papers show how Keynes, in public a mild man, was dispatched by Lloyd George,
14 ns the door to the manse. He is 65, a mild man, no fire and not much brimstone,
15 merited was not forthcoming. Clark, a mild-mannered man, has a bland public pers
```

16 Malcolm. For an essentially gentle, mild-mannered man, Malcolm is still inclin
17 g wicket-taker. Kumble is a modest, mild-mannered man. Off the field, he wears
18 te, but he is, first of all, a man of mild manners. If this man gave you an orde
19 suited and tied as Mr Robert Andrew, mild-mannered surveyor; a man doing well i

Figure 6.19 The concordance of ‘mild’ co-selected with ‘man’ in the 1995 *Times* newspaper corpus

Not only is ‘mild’ not very frequent when it describes people, but it also refers to a permanent personality trait, a habitually shown one. Expanded contexts showed that ‘mild’ may point to external appearances or behaviour only, as is also shown by the frequent adjective ‘mild-mannered’. In this light telling someone to be mild seems a little illogical. One may strive to be modest or resolute, but it is hard to see mildness as a character feature that is cultivated. Also, mildness seems to be a trait that exists in relation to other people or is shown in their presence. It is the opposite of violence, whether real or pretended, but it must exist in relation to others.

It follows from the above description of the behaviour of ‘mild’ in the *Times* corpus that the states of affairs it usually creates do not include the event described in the Pushkin’s line. Arguably, certain sacrifices may be made in order to retain the rhyme and the rhythm of the source verse in the translation. To this end, a slight shift of meaning may be tolerated. However, in our translation a shift of meaning is accompanied by an unnatural state of affairs that does not exist in the target language and is bound to ring false to a native ear.

This investigation would not have been complete if the Russian imperative *смирись* (‘resign yourself’) was not studied for contexts of situation. The search was limited to the literary subcorpus of the Russian National Corpus and yielded 68 lines, not counting repetitions. The interpretations given to the contexts were based on context clues and contexts of situation. The given classification was developed during the work with expanded individual contexts:

Table 6.1 Contexts of situation of the Russian imperative *смирись* (‘resign yourself’) in the literary subcorpus of the RNC

meaning of *смирись* in context	**number of lines (out of 68)**
a call for Christian submission/resignation	12
a call to submit to fate	10
a call to come to terms with a bad situation	10
a call to divest oneself of pride	10
a call to submit to a mightier force	10
a call to submit to authority	6
a call for inaction due to the impossibility of success	4
a call to calm down	4
a call to cease destructive action or a bad habit	2
a call to make peace with (your brother)	1

A call to be mild involves not showing violence to others or before them. The contexts of situation created by the Russian imperative do occasionally point at that meaning (viz. a call to submit to a mightier force, authority, give up a destructive action). Still, there are no ‘others’ in Pushkin’s verse. Indicatively, the imperative is mostly used to denote Christian resignation, then submitting to fate, coming to terms with a bad situation etc. It certainly points to an inner state of mind.

In sum, the case of this particular line showed inadequacies of translation attempted by a non-native speaker of the target language without consulting reference corpora. The translator’s difficulties arose from his unawareness of subtext created by grammatical strings and of states of affairs created by lexical collocates. So far, the suggested improvement of the grammar, through replacing ‘in’ by ‘on’ in the grammatical string of this verse, has had only moderate impact: diminishing the sense of urgency has not eliminated the sense of action and conflict.

6.3 Creating a corpus-based translation

This section will offer a corpus-based translation of the line which has so far been in the focus of the chapter. Moreover, as an afterthought, it will compare the corpus-based translation with an existing one published by a renowned scholar.

While criticizing an existing translation, it is only fair that the critic should offer a better one. Before this analysis, in particular of the word *смирись*, I did attempt several. Not belonging to the rank of experienced and renowned translators, but hopefully possessing some talent and understanding of the process, I came up with three dissatisfying versions. They contained the key words that I deemed the stanza needed to contain, and followed the relevant structure and the overall meaning. They did not contain approximate translations that deviated substantially from the original context(s), or that (according to myself) created wrong states of affairs. Nevertheless, they were far from Pushkin. The emphases seemed to shift from Pushkin's, the structures sounded unnatural, the final product just not inviting the amused and awed response it should. It just did not seem possible to keep Pushkin's meaning without violating the original rhythm and rhyme, or the other way around. My translations seemed, to me at least, fake. Here is one:

> Should life choose your hopes to cheat,
> Do not grieve or rage in vain:
> Come to terms with loss and pain
> Only days of joy to greet.

However, after my analysis of the contexts of situation of *смирись* ('resign'), yielded by the literary subcorpus of the RNC and classified in the previous section, I suddenly, in a wave of inspiration, wrote the following:

> In despondency, resign,
> In the day of joy believe.

And then the whole stanza wrote itself:

If life ever does deceive,
Don't oppose *the will divine*:
In despondency, resign,
In the day of joy believe.

Far from claiming that this translation is perfect (at the moment of writing I was still not aware of the one that, I assumed, had to exist and satisfy all criteria – namely, of the translation by James E. Falen), I nevertheless claim that it is better than the one I wrote before being introduced to the contexts of *смирись* (resign). My feeling that it has the naturalness of Pushkin's verse while retaining the meaning and spirit of the original is supported by the fact that I am capable of checking the line discussed in this section against my statistical table and proving that the meaning of the original has been retained. The previous line, containing 'will divine' (in italics), is in accordance with the contexts of *смирись*, in particular 1, 2 and 4. The key parallelism of the last lines is retained, although 'day' is gone from the key line discussed in this chapter. The other replaced part is the very subtext of 'in the * of' or 'on the * of', which may not be a bad idea given that in English even the subtext of the line beginning with 'on' might have proved too fight-evoking to be used in the English translation. I may only add that the line 'Don't oppose the will divine' was at the time of creating this translation in agreement with my native speaker's understanding of Pushkin's meaning, whose attitude to heavenly power, I felt, was precisely this: not exactly Christian but still aware of an unspecified divinity.

This feeling was confirmed by the corpus of his work. His prose corpus revealed the use of the word 'God' as mostly delexical, in expressions like 'Thank God' and 'God knows'. The degree of delexicalisation varied, but all these expressions were used by Pushkin to enliven the speech of the described characters, whether real or fictitious. In the rest of cases, 17 out of 229, the word appears in a quote. 'God's will' is used in Pushkin's prose corpus 6 times, delexically, once as *божия воля* and 5 times as *воля божия*. Three

times it is used by characters of novels, once by the persona of the narrator, and once by Pushkin himself, but the use remains delexical throughout. The 361 occurrence of the word 'god' in Pushkin's poetic corpus presents a diverse picture suggesting the need for a serious and thorough corpus study of what may have been deep-rooted and sincere atheism. In any case, whether the reference is delexical (in most cases), ironic, heretical (in some) or plural ('gods' are mentioned in 53 concordances out of 332), the overall impression is precisely what my intuition first suggested: Pushkin in his dealing with the divine did not go beyond some abstract notion of God, or a divine will, and even that he very rarely acknowledged.

I believe it is important that, as a native speaker of Russian, I was intuitively satisfied with the new translation, apart from the reliance on Russian corpus findings. So how did it come about? I maintain that, apart from conscious awareness of the contexts in which the imperative *смирись* ('resign yourself') is used in Russian literature, its 60-odd contexts in the RNC gave me the personal experience that I lacked before. I accumulated enough experience of the word to have become capable of an inspired translation, just as – almost just as – the poet must have been inspired at the moment of writing. The accumulated experience enabled me to express the source meaning naturally, not violating English subtext while creating naturally fitting states of affairs. My hypothesis had been that the translator would acknowledge the states of affairs created by the node and operate with them on a conscious level. What I did not expect was that the contexts of situation would penetrate the translator subconsciously – one may add, like the use of semantic prosodies. This favourably compares with a situation in which the translator is intuitively inspired to create a state of affairs which differs from the one in the source text. In my case, the corpus helped to create not only *informed* but also *inspired* translation.

6.3.1 States of affairs created by co-selection

I have objected to the use of the word 'grief' in the translation discussed in the previous section and preferred 'despondency'. This must also be verified. The *Times* corpus gave 546 lines containing 'grief'. This is the list of its collocates (grammar words have been omitted): 'come' (37), 'good' (23), 'came' (21), 'private' (16), 'anger' (13), 'shock' (11), 'death' (10), 'loss' (8), 'rage' (8), 'outpouring' (7). 'Come' is mostly part of 'come to

grief' (32 lines out of 37). The rest are 'have come through *acute* grief', 'come to *terrible* grief', ''the grief of *dead* comrades', 'repressed grief and *anger*'. 'Good' is part of 'good grief'. 'Private' is part of 'private grief', the contexts pointing to a possibility of intrusion, and thus suggesting that grief may not be private (whoever would talk of 'private despondency'?). 'Outpouring' is that of 'grief' in all the 7 cases.

The emotion of grief, then, is caused by death and loss and co-occurs with anger, shock and rage. All these collocates are parts of the noun phrase where 'grief' is present. Grief is also described as private, among other things in contexts in which 'intrude' or its derivatives appear as collocates (8 lines out of 16). This description is far from surprising and agrees, I believe, with anyone's intuition and accumulated experience of the word. What comes as a surprise, however, is that 'grief' and 'despondency' are not easily distinguished from one another on the basis of dictionary definitions alone. *Cobuild Dictionary* tells us that while 'grief is a feeling of extreme sadness', 'despondency is a strong feeling of unhappiness caused by difficulties which you feel you cannot overcome'.[28] While collocates in the examples given by this corpus-based dictionary throw better light on the difference in meaning, let us consider collocates of 'despondency' in detail.

If the *Times* corpus gave 546 lines featuring 'grief', those containing 'despondency' amount to 36 lines, which is a 15 times rarer occurrence. This alone tells us that despondency is not newspaper material: we assume it is not caused by events announced in newspapers. The concordance below shows that it does not set in abruptly, like grief, and that it is a protracted feeling, dull rather than intense. Its most frequent collocate (grammar words omitted) in the *Times* corpus is 'gloom' (5). Here is the full concordance:

```
1 om their anger.   After months of despondency about their dismal poll results, t
2 t and showed off his bowling. Any despondency about his political fortunes must
3  <Group>  WITH tales of gloom and despondency all around Olazabal's foot, Faldo'
4 , successes and failures, joy and despondency among the 40 players who will play
5 are apparently spreading fear and despondency among Conservative MPs.   The "lea
6  own worth and flunking it brings despondency. An inability to win the game in t
7 anically busy life and the tragic despondency and madness of his last years.  An
```

[28] The study was conducted in 2013. In 2019, Cobuild defines grief as 'very great sadness, especially at the death of someone' , and despondency as 'unhappy and with no hope or enthusiasm' (https://dictionary.cambridge.org/dictionary/english/). I am retaining this example as it is indicative of the usefulness of corpora generally. Besides, my analysis still contains elements not referred to in Cobuild.

```
 8 vertheless still manages to exude despondency and shyness.   Thus, after beginni
 9 You need only follow the trail of despondency and recrimination left in their wa
10 m or go on the dole."   There was despondency and pessimism at first among colle
11 ttering candle. The pose suggests despondency, and he hides his eyes behind a ha
12 income.   There is bitterness and despondency at the BFI's headquarters_in Steph
13 for a vanity press. Deepening his despondency at his fiction's lack of acclaim i
14 ll today, in an era of publishing despondency, booksellers swear that the green
15  far, they were cast into greater despondency by a performance ominously resonan
16 ly evaporated. Instead, an air of despondency has descended after four successiv
17 0, which Blake describes only as "despondency" his debts and his womanising, tog
18 how vernal depression, an unusual despondency in the spring. The popular explana
19  The Nikkei reached its trough of despondency in early July, but since then shar
20    <Story>     <Group>  GLOOM and despondency in the quoted property sector seem
21 mething. There is a great deal of despondency in the area because the police see
22 try's growing optimism changed to despondency inside 90 minutes.   "I am angry,"
23 s ago", thereby strking gloom and despondency into the heart of the would-be tra
24 nning himself added to the air of despondency last week. He told the Scottish Ca
25 alf lyrical. Only at the end does despondency prevail, in Miroslaw Balka's woebe
26 alf lyrical. Only at the end does despondency prevail, in Miroslaw Balka's woebe
27 alf lyrical. Only at the end does despondency prevail, in Miroslaw Balka's woebe
28  intensified the air of gloom and despondency settling over a sport that had tho
29  Stokes, shot himself in a fit of despondency.   The second, Robert Fitzroy, inv
30 to hospital could bring alarm and despondency to either camp, and talk by a vote
31 reading hither and yon; gloom and despondency were etched on the sponsors' faces
32 ble draw at Ibrox was followed by despondency, when Tommy Burns's side could col
33 nt as one more pace away from the despondency which almost destroyed his career.
34 erial employees showed widespread despondency, with many living in fear of losin
35 compound the sense of loneliness, despondency, worthlessness and despair.   By i
36 y braille," Stark said, with mock despondency. "You look round, hearing clonk, c
```

Figure 6.20 The concordance of ‘despondency’ in the 1995 *Times* newspaper corpus

Despondency is shown as a lasting emotion that strikes groups as well as individuals and whose causes lie in financial and professional failures or prospects thereof (as should be expected of a newspaper corpus), and not in shocking events like death. The emphasis is on duration, it seems, and not on the outrageousness of the cause. Except for line 29, it behaves like a persistent colouring, not like a sudden outburst. The emotion is not violent (perhaps with the exception of line 22 (‘I am angry’) and line 30, where it collocates with ‘alarm’), although it can be extremely destructive (lines 29, 33). ‘Gloom’ is the only frequent collocate.

Let us see to what degree the Russian word *унынuе* corresponds to ‘despondency’, which is its dictionary equivalent, and whether my claim that *унынuе* implies a passive state, and not an active one, is correct (on this claim I based my rejection of ‘grief’ as an equivalent, and also doubted the subtext in ‘in the * of’ and in ‘on the * of’). *Унынuе* in

Russian is a term used in the Russian Orthodox Church[29] as an equivalent to Greek *ἀκηδία* and Latin *acedia* – so much so that one is tempted to translate the line in question as '*In acedia*, resign'. *Acedia*, as we are aware, is a state of lethargy or spiritual indifference, certainly passive. The associations Pushkin or the reader was/is bound to have are certainly religious (especially in the context of *смирись*), but not only those. The word *уныние* has been widely used in worldly contexts, as the RNC newspaper corpus shows by yielding 437 occurrences. Therefore, the choice of *acedia* would not include occasions of despondency without religious connotation, which Pushkin must have had in mind, having been inclined to atheism and not to religious devotion, as we have seen from our brief search of the word 'god'. On the other hand, the chosen word 'despondency' encompasses general emotional states, as we have seen from the *Times* concordance, and leaves space for religious associations (think of John Bunyan's *Slough of Despond*).

6.3.2 Using parallel corpora when producing corpus-based translation

'Despondency', then, is a dictionary equivalent of *уныние*, informed with both worldly and religious meanings of the word. Still, it was useful to check how the words *уныние* and 'despondency' are generally translated into English and Russian respectively. For this purpose, I searched the English-Russian Parallel Corpus of the RNC. Such parallel corpora 'enable the researcher to study actual correspondences rather than to rely on translational equivalents found in dictionaries' (Ebelin 2013).

What English words are translated into Russian as *уныние*? Bearing in mind the diversity of sentence structure, I am listing all parts of speech as they cropped up in the 67 contexts, in order of frequency:

gloom, gloomy, gloomily	11
depressed, depression	6
unhappy, unhappiness	4

[29] Also in the Serbian Orthodox Church, as *униније*.

The rest include: ‘affliction’; ‘aggrieved’; ‘annoying’; ‘bleak’; ‘cast’ ‘down’; ‘consternation’; ‘crestfallen’; ‘dejection’; ‘**despondency’**; ‘demoralization’; ‘discouraged’ (2); ‘dismally’; ‘dismay’; ‘distress’, ‘distressed’; ‘downcast’; ‘downhearted’; ‘drag’; ‘dreary’, ‘dreariness’; ‘miserable’; ‘moodiness’; ‘pall’; ‘rueful’, ‘ruefulness’; ‘sorrow’; ‘**Slough of Despond’ (2)**; ‘the dumps’; ‘wretched’.

Уныние is translated into English as ‘depressed’, ‘depression’ 6 (out of 12), ‘dejected’, ‘dejection’ (3), also as ‘anguish’, ‘**despondency’**, ‘sadness’, and ‘melancholy’. ‘Despondency’ is translated as ‘уныние’ in 2 cases out of 7. Although this search gives legitimacy to my choice of ‘despondency’ as an equivalent to ‘*уныние*’, I was particularly struck by the fact that another legitimate equivalent to *‘уныние’* appears to be ‘gloom’. I was struck by it because during the search the following line occurred to me:

On a gloomy day, resign

While preserving the meaning of *уныние*, the adjective ‘gloomy’ made the introduction of ‘day’ possible, saving the parallelism present in the original:

On a gloomy day, resign,
In the day of joy believe.

The parallelism of ‘day’ is saved while the questionable subtext of ‘in the * of*’ or ‘on the * of *’ is avoided (including the tragic impression derived from lines such as ‘on the day of her death/the funeral’ etc.). [30] I found this version particularly satisfying because ‘a gloomy day’ reminded me of gloomy weather, and a spell of such weather is always followed by a sunny one, as in life a period of sadness is replaced by one of joy. Of course, the event in Pushkin is not measurable as a day. Its consequences will only pass away if one resigns oneself to fate.

[30] While it was later discovered that ‘on the * day’ revealed a similar semantic prosody to ‘on the * of *’, it was also noted that ‘on the day of *’ left an impression of momentousness that cannot be ascribed to Pushkin’s line.

6.3.3 Deciding on subtext in corpus-based translation

However, a new question arises – do we choose ‘on a gloomy day’ or ‘in a gloomy day’? And is it ‘a day’ or ‘the day’? Louw (2000) attaches key importance to the use of indefinite versus definite article. According to Louw, the use of the indefinite article implies unwelcome uncertainty (Louw 2000).

The following search lines were checked: ‘on a * day’, ‘on the * day’, ‘in a * day’ and ‘in the * day’. The line ‘on a * day’ is used both in positive and negative contexts in the *Times* corpus. ‘Weather’ words predominate in the wildcarded slot. This is the frequency list of these words in the 262 concordance lines:

‘weather’ words (‘sunny’, windy’ etc)	71
‘clear’	30
‘good’	29
‘bad’	18
‘season’ words (‘summer’, ‘wintry’, ‘June’ etc.)	18
‘particular’	7
‘single’	7
‘typical’	7
‘fine’	5
‘busy’	4
‘given’	3
‘calm’	3
‘nice’	2
‘normal’	2

Figure 6.21 The frequency list of the collocates in the lexical slot of the string ‘on a * day’, from the 1995 *Times* newspaper corpus

The search ‘on the * day’ yielded 1020 lines, but the wildcard proved to hide surprisingly unvarying collocates. They were mostly ‘last’ and ‘same’, ‘opening’, ‘final’,

and ordinal numbers, most often ‘first’, ‘second’, ‘third’ and ‘fourth’. Very few other collocates appeared. Those included ‘Christmas’, ‘penultimate’, ‘unfortunate’ mentioned once, and ‘big’ several times. As might be expected, the collocate ‘very’ (13 times) turned out to emphasise a coincidence, either fateful or ironic. This may have intuitively predicted simply from the general knowledge of the articles. ‘A day’ is one of the many, whether good or bad, gloomy or otherwise. ‘The day’ has a bearing on the situation in question. The difference is between a routine day and a special one. Therefore, ‘On a gloomy day, resign’ sounds more casual, and ‘On the gloomy day, resign’ will be more consequential.

Earlier, the preposition ‘in’, in the context of ‘in the * of *’, proved to be centred around more dynamic states of affairs, if ‘day’ fills the first lexical slot. The search of ‘in a * day’ revealed that the most usual adjective describing ‘day' was ‘single (17 occurrences in 50 lines, or one third):

```
 1 phone calls on one occasion 100 in a single day. After White's first transatlant
 2 uth, or on into Italy and Spain in a single day, are also at risk."   So, for an
 3 umbed to frostbite and exposure in a single day during the fighting and three we
 4 rase a prehistoric field system in a single day: gratuitous loss is terrifyingly
 5 attle of Waterloo, which he won in a single day. He came top in Country Life mag
 6 than 3 per cent, a big movement in a single day. In London, the FTSE 100 index f
 7  This strange novel takes place in a single day in Portugal, during which the na
 8 pan, where he lost $600 million in a single day last year when a punt on the yen
 9 are price climb from $28 to $75 in a single day, leaving James Clark, its chairm
10 s that you can get every season in a single day," Mr Tait said. "If you get a re
11 ver the mountains to Jarandilla in a single day, reclining in a curious leather-
12 est quantity of petrol supplied in a single day since the advance into Germany w
13  of The Daily Telegraph to 30p. In a single day that wiped 35% off the company's
14  and mortar rounds hit the city in a single day.   The Serbs see the peacekeeper
15 the value of electricity shares in a single day when Professor Stephen Littlechi
16 zze of cultures and traditions. In a single day you can see a Roman site, a Byza
17 he authentic Arctic wilderness. In a single day you can drive through a dozen di
```

Figure 6.22 The concordance of ‘in a single day’ from the 1995 *Times* newspaper corpus

As we have seen, the collocate ‘single’ implies an inordinate amount of change, welcome or unwelcome, in one day. ‘Full’ (3) is the only collocate producing a negative context. The other collocates are create positive states of affairs, including those one would not expect to behave this way: ‘superb’, ‘successful’, ‘rewarding’, ‘Valentine’s’ (2), ‘fair’,

‘hard’, ‘distant’, ‘punishing’, ‘normal’. Therefore, due to the positive prosody implying inordinate progress this pattern is unsuitable for the line we are translating.

The concordance of ‘in the * day’ is not wholly positive and, at first sight, difficult to make sense of. Semantic prosody aside, there remains the fact that ‘in the gloomy day’, if chosen, would preserve the parallelism with the next line (‘In the day of joy believe’). The alternative being ‘on the gloomy day’, it matters very much what I will choose. The lines in which the preposition ‘in’ belongs to an irrelevant structure, e.g. ‘culminating in’, have been eliminated so as to avoid confusion, leaving 20 lines:

3 fence, was **treated less seriously** in the present day than in the past. However,
5 **shells across Sarajevo yesterday, in the worst day of fighting** since Nato threat
6 ich was **so radical and innovative** in the Bard's day become so again, but now as
7 gine sometimes perhaps **too fondly** in the present day their echo at Lord's: Lik
13 t work you do, it can be **crucial**. In the modern day you **wouldn't really be able**
14 **to be remodelled without surgery in the first day or two after delivery**. <
15 ian **casualties continued to mount in the fifth day of fighting for control of** th
19 , in **resetting the original novel** in the present day with a corresponding invers
21 **sessment of the composer's status** in the present day as with market consideratio
29 INCENZI'S **Forbidden Places starts** in the present day, but a mere five pages in w
38 o be the Boy Scouts' **headquarters** in **the Shah's day**. There is **deep hatred_**betw
42 e the same tourists **several times in the same day** in the morning at the shops, a
43 **market, you have to put an offer** in the next day in case the price goes up," sa
44 r **salary cheque is due to be paid in the next day**. It is the **final straw** and you
45 d of attack as "terrifying". **When in the same day of the third Test match** at Ade
46 ow set **to move across the country in the next day** or so. As many people are alre
47 rk **central defender**, John Hughes, **in the next day or two**. A presumably modest fe
48 now that **the selectors will meet in the next day or so to confirm that**, in Nico
49 junk-yard dogs look like poodles. **In the past day or two he has been_**all gap-too
50 n, said: **"I expect I will see him in the next day or so. It is a private matter**

Figure 6.23 The concordance of ‘in the * day’ from the 1995 *Times* newspaper corpus

Judging by the concordance of ‘in the * day’, the preposition ‘in’ is indicative of movement, action and change. Besides, while the concordance of ‘in the * day’ consists of 20 lines, the one of ‘on the * day’ lists 1020 lines. The frequency data speaks in favour of using the preposition ‘on’, as this usage is more routine.

Therefore, this is the corpus-based translation I arrived at (where, strictly speaking, only lines 2 and 3 are corpus-based):

If life ever does deceive,
Don't oppose the will divine,
On the gloomy day – resign,
In the day of joy believe.

In the following section this version will be compared to an existing translation, offered by an expert and a native speaker.

6.3.4 Corpus-based translation versus an existing professional translation

I consider it fortunate that I encountered James E. Falen's translation of Pushkin's stanza only after I produced my own corpus-derived one. Given the American translator's renown, I may not otherwise have had the patience or the courage to attempt the long process of corpus analysis before looking at Falen's work (Pushkin 2009: 86):

Friend, should life in time deceive you,
Do not sorrow, do not fret!
Humbly bear the days that grieve you;
Days of joy await you yet.

By the future live our hearts;
Though the present may be drear –
All is fleeting, all departs;
What has passed will then be dear.

At first sight, to me this seemed to be a winner. The literal 'resign' has been replaced by the paraphrase 'humbly bear', which retains the meaning perfectly without losing any of the naturalness of the original. This native speaker's advantage is also reflected in the use of

the verb ‘sorrow’, whose existence did not spring to mind in my case. The filler vocative ‘Friend’ fits perfectly, creating the tone of the message that corresponds to the original. The charm and ease of the verses is certainly Pushkinian (my intuition was saying), and by using paraphrase (e.g. ‘humbly bear’, ‘the days that grieve you’) the author has successfully avoided the all too imminent (and all too familiar to me) danger of the expressed sentiments sounding like a list of clichés.

Still, there remains something to be said in favour of the corpus-based translation. The rhyming pattern Falen chooses is abab, whereas in the original it is abba. The manoeuvre is understandable and, of course, allowed. However, it renders the overall tone lighter, more good-natured, less pensive than the original rhyme abba would have done.

In addition to the change of abba to abab, the ‘b’ rhymes are ‘fret’ and ‘yet’. Although the metre has been retained, the change of the rhyme scheme makes the last line sound shorter than the original and end in a stressed syllable. On the other hand, in Pushkin, the concluding verse is one syllable longer and the syllable is unstressed, making it sound more contemplative (after all, that is why this rhyme is called feminine). The alterations in the rhyme scheme of the translation relevantly change its tone, making the rhythm almost dancing. How easy it seems, upon reading the translated stanza, to live through yet another bad patch, and how certain we must feel of our fortune changing for the better eventually.

The choice of words also contributes to the light-heartedness of tone in Falen’s translation. ‘Fret’ appears in the second line ‘Do not sorrow, do not fret’. Let us look at those contexts of ‘fret’ in the *Times* reference corpus which are preceded by ‘not’:

```
10 e Rugby World Cup leave you confused? Fret no more, ROLAND WHITE's simple guide
12  ight "jumbulances" in the City, don't fret that someone has hogged all the parki
13 ahs in southwest England.   But don't fret, help is at hand with Garden Shield,
14 acts cancer. What will they do? Don't fret: there are no difficult ethical decis
15 ry as a Millennium project. But don't fret: it's pretty good inside.   Watch out
26 lleged to have said that he would not fret "if a policeman is killed because he
27 Charles addressed.   Charles need not fret about us embracing alien ideology and
28 ilms.   However, aficionados need not fret too much about Bond's change in alleg
29 death from a broken heart. But do not fret, mother and son will be re-united abo
30 cause he is behaving badly I will not fret."      </Group>   </Story> </Article>
```

```
31 ary goes absent without leave, do not fret. She could be at London's Olympia 2 t
32 cause he is behaving badly I will not fret."   Mr Narayan, 57, also handed out l
```

Figure 6.24 The concordance of 'fret' preceded by 'not' in the 1995 *Times* newspaper corpus

The use of 'fret' seems to range from the jocular to the plainly ironic. The use seems to mean (bar the sarcasm in line 26), do not fret, all your worries are in fact trivial and unworthy of serious concern. The definition of the verb 'fret' in the *Longman Dictionary* supports the *Times* usage: 'to feel worried by small and unimportant things'.[31] The verb 'fret' certainly should not be co-selected with the verb 'grieve', appearing in the next stanza (the noun 'grief' in the *Times* has been discussed earlier). The line would have been better translated as 'Do not grieve and do not rage', as the anger (not worry, as implied by using 'fret') is directed against the circumstances, fate, God, 'deceiving' life (as in 'Rage, rage against the dying of the light' by D. Thomas). It is plain that the co-selection of 'fret' with 'sorrow', and with 'grieve' in the next stanza, is metre-induced.

'Days that grieve you' also arouse suspicion. I have already shown the difference between 'grief' and 'despondency'. It is plain that Pushkin refers to days of depression rather than event-induced grief, and that life's deception does not amount to bereavement, for example – an event that cannot be erased by a turn of the Fortune's wheel, as despondency-inducing events may be. Matters are not helped by 'do not sorrow' in the second line. This verb is much rarer than the corresponding noun, and more frequent in COCA than in the BNC. The contexts studied point to a feeling of deep and protracted sadness caused by a loss, not necessarily death. Very many contexts of the noun in COCA philosophically suggest that a state of sorrow is in time replaced by that of joy, or that without the former we could not experience the latter. However, the word is not acceptable in Pushkin's context because the corpus clearly shows that sorrow is sadness of such depth

[31] COCA creates a somewhat different picture: apparently, in American English the word is stronger, and the contexts less trivial. However, they are still too trivial to justify the use of the word. Moreover, 'to fret' means 'to worry' and not 'to be angry' as in the original. Finally, the change of tone induced by the combination of the abab versification pattern and by the rhyming of 'fret' with 'yet' cannot be disputed.

that it cannot be missed later in life, and this is literally the message of Pushkin's second stanza.

It is obvious that a very fine balance needs to be struck in the translation. The context of situation that was to be created was in fact somewhere in between 'fret' (too light-hearted) and 'grieve' (too serious). In this light, Falen's translation, being, of course, intuitive, might seem like a string of almost-acceptable choices that seem more doubtful after referring to corpora than on the first reading. On the other hand, the corpus-based translation does not contain a choice that is not justifiable and scientifically based. Moreover, its choices are procedurally governed: starting with vocabulary and letting the grammar chunk itself is no longer the only option. The research of grammatical strings and their lexical collocates may significantly contribute to the quality of translation.

Falen's intuitive translation has enormous merit, and the reader might prefer it to the corpus-derived one offered here. However, I would like to argue that such a preference would stem not only from the charm and naturalness of Falen's version, but also from the overall lighter tone which does not exist to the same degree in the original.

Speaking of the deficiencies of Falen's translation, it is fair to dwell on those in the corpus-derived one. The one that first springs to mind is the first line 'If life ever does deceive'. The first two words defy the chosen trochee pattern (we have seen that Falen's trochees are remarkably regular). Although from then on the rhythm flows smoothly, matters are not helped by the filler 'does' which both keeps the metre and the rhyme deceive/believe. It is my educated guess that no English poet would have ever been inspired to start a poem with 'If life ever does deceive', not even in the 1820s; it would more likely have been 'If you are deceived by life' (bringing the question of rhyme in the foreground – strife? knife? survive? revive?)[32] Also, the filler 'ever' does not exist in the original and thus may be lessening the implied possibility of life's deception (Falen's 'in time' is also a filler but has different implications). The consolation is that this particular line is not corpus-derived at all, as in this chapter the focus was on the third line, which forced me to consider the second and the fourth.

[32] I may be too harsh on my translation. Shakespeare allowed himself to write 'Rough winds do shake the darling buds of may'; whether this filler was still acceptable in the 1820s, Pushkin's time, ought to be checked in a diachronic or poetic corpus.

Then, it is fair to mention that the second line, ‘Don’t oppose the will divine!’ covers *не сердись* (‘do not be angry’), but not *не печалься*, translated by Falen’s ‘do not sorrow’. It might be argued, however, that in my translation sadness is understood, as it is demanded by the context of situation. On a gloomy day one is sad by inference (that is why the day is gloomy), but not everyone is also angry.

As for the last line, it is also not corpus-derived, but it does make use of ‘in the * of’, which was discussed in this chapter and whose implications are welcome at least in part (constructive activity we do need, but elements of fight-picking we do not). ‘In the day of joy believe’ (which is what Pushkin is saying in the original) may sound less optimistic than ‘Days of joy await you yet’. This could partly be due to the plural of days (days of joy sound better than one day of joy, although ‘day’ is singular in Pushkin), partly to the effect of optimism communicated by ‘yet’, and partly to the semantic difference between the hopeful ‘believe’ and the confirmatory ‘await’. Also, the ‘/eit/’ in ‘await’ happily resonates with ‘yet’ and ‘fret’).

Finally, if in Pushkin the rhyming pattern is feminine-masculine-masculine-feminine, in the corpus-derived translation the rhymes are all masculine, which may contribute to an air of finality and determination. Perhaps, choosing 'on **a** gloomy day' would have been wiser because it does not have an air of relentless finality. All in all, while the corpus-derived translation arrived at the necessary state of affairs by careful exploration of corpus contexts, Falen reached a similar overall impression by counterbalancing lexis that is too strong (‘sorrow’, ‘grieve’), with words that are too light (‘fret’, ‘yet’). While Pushkin’s advice is to believe in the day of joy, Falen is much more confident. The impact of ‘sorrow’ and ‘grieve’ is further overridden by a more light-hearted rhyme scheme. If we involved actual respondents and asked them to choose the better translation, we might receive feedback based on the overall tone rather than on the degree of linguistic equivalence. Nevertheless, the corpus-based translation may be further improved by the same principle being applied to its other lines.

6.3.5 Concluding remarks on sections 6.2 and 6.3

Section 6.2 intended to analyse the third line of an existing translation of a stanza by Alexander Pushkin. My search of 'in the * of' yielded a picture of action verging on conflict, which made me wonder whether this was not an unacceptable grammatical pattern to express Pushkin's call for resignation. When consulting a similar search line - 'on the * of' – I understood that this line's prosody was also argumentative. This may have to do with the fact that the consulted corpus was that of the *Times* newspaper, whose topics must of necessity be wars, sports, legal cases and generally conflict-based issues. However, the main factor distinguishing between the two lines turned out to be the quasi-propositional variables of the lines. The quasi-propositional variable 'day', appearing in the line 'on the * of', did appear as the QPV of 'in the * of'. The only corpus where it was on such a list was Google-Books – UK, where the contexts were mainly of the Last Judgement, and also once of battle and once of distress.

The conclusion so far being that the preposition 'on' needs to replace that of 'in' in the target line, I moved on to the search of the Russian imperative *смирись* ('resign'). It yielded contexts that inspired me to re-create the previous line in English as 'Don't oppose the will divine', instead of the more literal 'do not grieve and do not rage', for example. The need to re-create the previous line was brought on by earlier unsuccessful attempts to translate the poem and at the same time keep Pushkin's versification and naturalness of expression. This translation of the previous line inspired the whole stanza, where the target line ran 'In despondency, resign'.

After I had gathered evidence showing that 'despondency' is more adequate than 'grief' as a translation of the Russian *уныние,* I found a better alternative for the noun phrase *день уныния* ('gloomy[33] day'). This discovery was helped by consulting the RNC's

[33] The term *gloomy* offers an almost *a priori* connection between the weather and the elements and human moods on the other. For this reason, the term shows only very slight delexicalisation into other meanings and is a powerful choice in the act of translation. Because gloomy people are often impassive, it is a term that almost *speaks* on their behalf, as we see here (Bill Louw, personal communication)

parallel corpus. It turned out that *унынuе* is the word that translators often choose as the equivalent for the English ‘gloom’ and ‘gloomy’.

The final step was to decide on the preposition and the article – whether it should be ‘in’ or ‘on’ and ‘a gloomy day’ or ‘the gloomy day’. It turned out that the preposition ‘in’, as in the case of the search line ‘in the * of’ suggested movement and change, and was 40 times rarer than the preposition ‘on’. ‘The’ was chosen because ‘a’ creates more casual states of affairs, while ‘the’ creates more consequential ones. ‘A’ is a legitimate alternative, but suggestive of the mundane rather than the significant. Also, ‘a’ attracts weather-related lexical variables, unlike ‘the’, and in combination with the adjective ‘gloomy’, it could suggest wrong associations.

The stanza I arrived at was a consequence of corpus search combined with the translator’s inspiration. Corpus verification affords a thorough scientific basis for all the choices made. The corpus research itself gave the translator experiential basis for poetic inspiration, which ensured a no less vital component – naturalness. Section 6.3 compared the corpus-based translation to the existing translation by James E. Falen, whose work retains the original metre, but alters Pushkin’s rhyme scheme, to the effect that it creates a lighter mood than in the original. The rhymes Falen chooses contribute to this mood. The lexical choices he makes may be found less than perfect when we check the states of affairs they produce in the corpus, e.g. ‘fret’ and ‘grief’. Compared to Falen’s translation, I kept the rhyming pattern but changed the feminine rhymes into masculine, which may be aggravating Pushkinian tone even further. Overall it must be noted that the corpus-based stanza could be at its faultiest where it is, in fact, intuitive (line 1 was never corpus-derived).

The importance of naturalness in translation cannot be over-emphasised. Section 6.3 has shown how it can be attained through corpus-based experience. However, it does not consider another factor that may help the translator’s inspiration. Namely, it was taken for granted in this chapter that the metre of the original must be retained. It need not be so. A translator may stick to the pattern of, say, the first (or not necessarily the first) line *as it naturally sounds in the target language*. For example, Pushkin’s first line, translated

without respecting the metrical pattern of the original, says in English: ‘If life ever deceives you’. This line may serve as a foundation for, say, the following final product:

If life ever deceives you,
Don’t oppose will divine:
In misfortune, resign,
Until misery leaves you.

Life is usually glum,
Happiness always nascent.
All will go, more will come,
And be missed in the present.

The principle adopted here is that the translator works from the first line (or any key line that springs to mind) as it is naturally rendered in the target language. The rest of the poem flows (corpus-assisted) in the metre of the first line. This is redolent of the authentic creative process when the poet invents the first line (not necessarily marking the beginning of the final version of the future poem) and works from it, adapting the actual expression’s versification accordingly.

It may be argued that trochees carry a different emotional suggestion for the recipient, perhaps a happier, more contented mood. This opens the question of the implication of different metres for different languages.

Obviously, the best way seems to combine the native translator’s naturalness of expression with the information gained from the reference corpus. Naturalness is a result of inspired as opposed to forced writing, of expressing content that is one’s own (authors do that) rather than another’s (translators do that). Still, although the lexical choices of the corpus-derived translation are scientific and justified, it is my intuitive impression that Falen gracefully and apparently without any effort occupies the golden mean between the cliché and the forced (unnatural) expression (although he does co-select ‘sorrow’ and ‘fret’, rhymes ‘fret’ with the optimistic ‘yet’ and changes the rhyme scheme in the direction of the

jovial). What is it in the form, as opposed to content or meaning, that makes all the difference between a cliché or forced phrasing and a work of art?[34] The issues of naturalness and inspiration will be discussed in Section 9.

6.4 Examining Falen's translation of *передо мной явилась ты*

Since the subtext of Pushkin's line *передо мной явилась ты* was extracted in the course of Section 5, this proves an excellent opportunity to compare the subtext in the original, so laboriously obtained, to that of the translated text. This is exceptionally useful given the remarkable gift of the translator and the difficulties he seems to have encountered when dealing with Pushkin's poetry:

In his *Note on the Translations* to Pushkin's *Selected Lyric Poetry* Falen writes: 'Pushkin's work is notoriously elusive for the translator and on many occasions as I worked on his poems, I had to abandon my efforts, finding many of the lyrics utterly untransferable into English, at least by me' (Pushkin 2011: xxii). It is all the more interesting to investigate the poems whose translations by Falen met his own very high standards and found their way into print. Falen's translation of Pushkin's line *передо мной явилась ты*, which was discussed in the final section of the previous chapter, intuitively incorporates much of Pushkin's subtext. This is Pushkin's original juxtaposed with Falen's translation (Pushkin 2011: 89):

Я помню чудное мгновенье:	I still recall a wondrous vision:
Передо мной явилась ты,	**That day when I beheld your face**,
Как мимолетное виденье,	A fleeting moment's apparition
Как гений чистой красоты.	Of perfect beauty and of grace.

The translator has kept Pushkin's rhythm and rhyming pattern, which, of course, restricted his lexical choices. The grammatical string he opted for in the second line does not contain either the verb 'appear' or the preposition 'before'; instead, his chosen string is 'that * when I * your *'. The search in the BNC yielded no matches, while COCA gave 6

[34] I owe this phrasing of the question to Professor Boris Hlebec, University of Belgrade.

concordance lines, all but one occurring only once. The string of words that occurred more than once – three times – was 'that's when I saw your'. Strictly speaking, this is a different grammatical string from the one used by Falen, as it has a grammar word in place of the first wildcarded slot, and I would have probably ignored it had I been given more material to work with. What made me stop at this string of words was that all the three contexts seemed morbid: two involved death and one described the bullying of a younger child by an older one, which, judging by the context, could have ended in suffocation.

Google Books – Fiction, in the 10 available contexts, death is a present motif in 4, arson in one, theft in one, adultery in one. The string is a transition to rescue or safety in 6 contexts; it is a transition to discovering a morbid issue in five contexts; to adultery in one, to romance (presumably) in one. There were 7 contexts of 'that is when I saw your'. Death is present in 2 contexts, romance in 2, birth in 2, rape in 2, incest in one. In 4 contexts the string is a transition to rescue, a saving moment.

The shorter string, 'that is when I saw', will be of interest mainly because of the consistent state of affairs. Out of 20 contexts, violence, imminent danger or death are present in 11 contexts (4, 5, 6, 11, 13, 14, 16, 17, 18, 19, 20); faith and religion in 5 (1, 2, 3, 7, 10); romance (wedding) in one (incidentally, the book is called *I Married a Preacher with Deep Dark Secrets*).

The string 'that was when I saw your' yielded 2 contexts in the Google Books – Fiction corpus: one describing the grief of a man for the disappeared wife and the dead child, and one of a murder.

Next, since the Google Books corpus does not process strings of more than 5 words, I looked at the contexts of those of the first 100 strings yielded by the searchline 'that * when I *' that could be – and were – followed by 'your' in the corpus. The strings arrived at by the search of 'that * when I * your' yield two lists of quasi-propositional variables. The first list usually denotes a moment in the past, either through 'is' or was' or through lexis like 'night', 'moment' or 'happened'. In the place of the second wildcard there was found the verb 'to be' as 'is' or 'was', verbs of cognition or perception ('saw', 'heard', 'knew', 'realized', 'met', 'found'). The states of affairs created here are those of significant transitions in the past, many involving morbid or romantic contexts. Romance is more

present in the strings that do not contain 'is' or 'was' in the first wildcarded slot and the lexical word 'saw' in the second slot; death, morbidity or danger are present less. Religious contexts found in the contexts of 'that is when I saw' were not found here.

On the whole it would seem that the subtext of Falen's line 'That day when I beheld your face' corresponds to Pushkin's subtext in *Передо мной явилась ты* in its evaluative ambiguity. In Pushkin the woman is liable to censure which is prevented by lexical collocates. Falen keeps all the lexical collocates implying transience and the supernatural nature of the vision, while his subtext in the language has been used either as a romantic transition or a dangerous one, and always significant. This is often a moment of rescue, but sometimes of realization of irreparable harm. In the contexts of romance and man-woman contexts generally there is no implication of the woman's guilt, or unlawful sexual provocation – Falen's verb does not have a grammatical feminine ending, but would it have changed anything? Crucially, if the string contains "saw" – a superordinate of "beheld" – the prospect of danger rises sharply, but the potential of saving the situation or salvation through faith also appears. However, "saw" does not necessarily chunk the same states of affairs as "beheld". What is the contribution of "beheld" to the quality of Falen's translation? In the first twenty contexts yielded by 'beheld your' in Google Books – Fiction 'beheld your' is used in elevated discourse, in older texts or their translations. Religion is the topic of 9 contexts out of 20, love of 6 contexts (with morbid elements in 2). In this, 'beheld' resonates with contexts of religious worship in the Russian subtext of Pushkin's line, as well as with contexts where the woman appears before the man carrying an aura of sexual attractiveness.

Generally speaking, Falen's translation of Pushkin's line does transfer many of Pushkin's subtextual implications into English, most of all through the very ambiguity of Falen's chosen grammatical string – what is not love, could turn out to be death. 'Beheld' co-selected with 'your' contributes to the ambiguity by adding elements of religious discourse, also present in Pushkin's subtext, and strengthening the element of romance, which becomes complicated by unrequited love, torture, morbidity, the absence of the wife or jealousy – indeed, out of the six love contexts only in two do we witness no complications but the usual one of wooing.

What Falen's subtext does not involve at all is the implication that the woman may be at fault. Love in Falen's subtext is often the tale of the first meeting, sometimes of disappointments, complications, partings and even rape – but the woman is not the one whose appearing, despite her attractiveness, is thought wrong or inappropriate. If in Pushkin's subtext the appearance of an attractive woman carries something wrong with it that we could read either as guilt or as provocation, what takes place in Falen's subtextual contexts has no such rule. Thus, Falen's subtext could be said to be as ambiguous as Pushkin's, but - more adoring.[35]

What subtextual implications would the translation carry if the verse had been translated more literally, for example, 'when you appeared before my eyes'? The use of the verb 'appeared' would have ruled out the choice of the noun 'apparition' in the next verse, but I am now focusing on subtext alone. Would any overtones of guilt or sexual provocation have been transferred into English? What about the overall negativity overridden by the positive tone and states of affairs created by the lexical collocates?

As for the overall negativity, in a line featuring 'when you appeared before' it would be preserved. This claim is founded on the quasi-propositional variables of the searchline '*ed before' in the BNC and COCA. The three most frequent QPVs in the BNC are 'appeared' (298 lines). 'happened' (160 lines) and 'died' (116 lines); in COCA they are 'happened' (994 lines), died (579 lines), testified (547 lines), passed (477 lines), appeared (468 lines). 'Happened' is a negative semantic prosody established in Sinclair (2003: 117-125). As for 'appeared before', the concordances below are yielded by the BNC:

1 first time and could not have **appeared before**. Miller's Vinca rosea (now Catharanthus rose
2 shall and his CEGB colleagues **appeared before** the 1984 Commons Environment Committee they
3 ter, 46, and Robert James, 31, **appeared before** Justice Day. Under his recommendation, the
4 ecret of the clue, and one day **appeared before** her with a dagger and a cup of poison, and
5 and Queen. They had regularly **appeared before** the rich and famous but this audience inclu
6 arly as 1476 a London priest **appeared before** the Ecclesiastical Court charged with showing
7 entre # Dr Jonathan Winterton **appeared before** the House of Commons Select Committee on the
8 tment in Singapore. He duly **appeared before** three or four venerable gentlemen who lectured

[35] This example shows the potential for the approach in other fields of research, like comparative literature and sociolinguistics.

9 aniels and Janet Rogers had **appeared before** a Sports Council' panel' and answered many que
10 heir eyes when these idols **appeared before** them -- on roller skates! Much laughter and ap

Figure 6.25 The concordance of 'appeared before' in the BNC

…and by COCA:

1 der of North Korea Kim Jong-un **appeared before** his people trying to gather support for his
2 when the CEO of JPMorgan Chase **appeared before** Congress today he had a lot of explaining t
3 of one of the crablike aliens **appeared before** us. We stared. # " A virtual ambassador? "
4 e novel base word consistently **appeared before** its related derivative, the target word. Th
5 get us there. CAMERON: When Lew **appeared before** the GOP-controlled House Budget Committee,
6 neral commanding the Cairo area **appeared before** protesters and told them, " All your deman
7 cision that based upon all that **appeared before** them, that the ability to lead Penn State
8 rry's mother, sister and cousin **appeared before** the committee. ROBERT-HEYER-COUS: We hope
9 theater, benevolent ghosts who **appeared before** performances to bless and encourage the ac
10 ked up, startled, when a paper **appeared before** her, in Mason's hands. She took it, and he
11 ter a ride, an ethereal figure **appeared before** me. He had a helmet with a visor, a hydrat
12 of May, Marion and his brigade **appeared before** Georgetown and started digging siege trenc
13 oncern for Emory's Marino, who **appeared before** Congress last April to challenge the educa
14 ltant who often works with him **appeared before** the Board of Adjustment or the Planning Co
15 re than 40 cases where Fleming **appeared before** the planning commission and county commiss
16 how the holy, ancient temples **appeared before** they were chipped away by art thieves. Clo
17 e to the 35. # When Guskiewicz **appeared before** the NFL's competition committee in Februar
18 he department's leader when he **appeared before** hundreds of people in the Bayview on Wedne
19 d gifts from the litigants who **appeared before** him. Nevertheless, in those days if I'd ne
20 or Puerto Rican independence, **appeared before** the U.S. Parole Commission in Terre Haute,

Figure 6.26 The concordance of 'appeared before' in COCA

Apart from the decidedly prevalent prosody of appearing before courts and committees, this lexical collocation creates states of affairs similar to Russian ones that are chunked by *перед* (‘before’) and *явиться* (‘appear’) (see Section 5.4.2). The power disbalance is there, as well as supernatural creatures – ghosts (COCA line 9) and ethereal figures (COCA line 11). God will appear before one in religious texts (not present in the given concordances). A powerful woman kills her rival (BNC line 4) – incidentally, is it possible that appearing women are prosodically at fault in English? There are nine concordance lines containing 'appeared before' from fictional prose in the BNC, on page 3 of the search. In four there is an appearing woman:

1. After her coat was thrown down on to the couch, to be followed by the long mud-fringed skirt and tattered voluminous blouse, there **appeared before** the child a fat woman, a very fat woman, in what seemed to be a clean blue-striped blouse and a long grey skirt with a fringe.

2. He lay on the goose feathers, looking at the beams, feeling the soft pelting rain massage away the binding layers of ambition, lust, wickedness, pain, revenge... This was where he could belong... And Mary as he had first seen her **appeared before** him again like an apparition as he slid slowly into a long sleep.

3. Not being able to find an immediate answer to these questions, Cassie crept towards the mirror and stared in. Her reflection, pale and tired but quite solid looking, **appeared before** her in a perfectly normal manner.' Jeez!' she whispered, in surprise.' So I'm really here; here in Johnny's bloody awful hall!'

4. There was a slight noise upstairs, then Tess **appeared before** his eyes. She did not see him, and stretched one arm up above her head. She yawned like a cat and he saw the red inside of her mouth. Her whole soul breathed out physical beauty.

Judging by these four examples, an appearing woman in English is faulty unless she is attractive, which is at odds with Russian, where attractiveness is relevant but does not interfere with the feeling that there is something wrong with either the woman or her appearance (see Section 5.4.2). The subtext of either guilt or sexual provocation persists in grammar strings of both languages – the guilt in English is not of an inferior kind, however – but as for states of affairs chunked by the verb and the preposition, oddity and attraction do not mingle. Therefore, it seems to me, on the basis of these limited samples, that Falen's subtext, in its aura of fatal outcomes, might be more convincingly ambiguous than the one created by '*ed before'.

6.5 Future directions

This section has investigated the potential of Contextual Prosodic Theory to evaluate existing translations and produce better ones on the basis of reference corpus data. The procedures deployed at various points in the course of the chapter may be generalised into a list of steps that future translators may wish to go through in order to ensure that meaning in their translation corresponds to the original. To sum up, in the future, with much bigger reference corpora, translators will be able to take the following steps:

1. Examine the poem and break it up into a string of language events.
2. Examine a given language event's subtext.

3. Examine every language event's lexical collocates and semantic prosodies.
4. Look for authorial specificities of usage in the authorial corpus.
5. Use lexical collocation to summon similar states of affairs in the reference corpus of the source language.
6. Study the states of affairs for the context of situation.
7. Call up similar language events in the reference corpus of the target language.
8. Study the lexis in these.
9. Strive to achieve a line of best fit in translation.
10. Check the result for subtextual implications and semantic clashes.

This checklist, however, is not the only implication of the findings of the chapter. Bearing in mind the moderate (despite its empirical foundation) success of the corpus-based translation, and the arguably more than moderate success of the translation offered by a gifted professional translator, it seems best to entrust the task of translating to a native speaker of the target language who is also a corpus expert. In this way, a line of best fit may truly be achieved.

7 The interaction of semantic auras in a poetic text

7.1 Introduction

This chapter, based on Louw and Milojkovic (2014), Milojkovic (2017a) and Милойкович (in press 2019b), illustrates in detail how semantic auras of lexicogrammatical collocation, and in particular grammatical strings, interact within a poetic text and in this way confirm one another's subtext. It is assumed that the advantage of studying a short poem is the great attention that a celebrated poet is expected to pay to the expression of his or her meaning. The chapter will also dwell on rhetorical devices, such as metaphor, pun and personification, that have emerged as a result of a poem's subtextual investigation.

7.2. 'The Circus Animals' Desertion'

In the previous sections, semantic auras in texts, whether original or translated, were discussed in their context in the poem. No attempt was made to study other lines in the same poem subtextually, except in Section 4.5. This approach is adopted in this chapter, where a poem will be considered as an independent whole in order to see how its elements are interrelated at the level of collocation. The text of the chosen poem by W. B. Yeats is given below.

The Circus Animals' Desertion

I

I sought a theme and sought for it in vain,
I sought it daily for six weeks or so.
Maybe at last, being but a broken man,
I must be satisfied with my heart, although
Winter and summer till old age began
My circus animals were all on show,
Those stilted boys, that burnished chariot,
Lion and woman and the Lord knows what.

II

What can I but enumerate old themes,
First that sea-rider Oisin led by the nose
Through three enchanted islands, allegorical dreams,
Vain gaiety, vain battle, vain repose,
Themes of the embittered heart, or so it seems,
That might adorn old songs or courtly shows;
But what cared I that set him on to ride,
I, starved for the bosom of his faery bride.

And then a counter-truth filled out its play,
'The Countess Cathleen' was the name I gave it;
She, pity-crazed, had given her soul away,
But masterful Heaven had intervened to save it.
I thought my dear must her own soul destroy
So did fanaticism and hate enslave it,
And this brought forth a dream and soon enough
This dream itself had all my thought and love.

And when the Fool and Blind Man stole the bread
Cuchulain fought the ungovernable sea;
Heart-mysteries there, and yet when all is said
It was the dream itself enchanted me:
Character isolated by a deed
To engross the present and dominate memory.
Players and painted stage took all my love,
And not those things that they were emblems of.

III

Those masterful images because complete
Grew in pure mind, but out of what began?
A mound of refuse or the sweepings of a street,

Old kettles, old bottles, and a broken can,
Old iron, old bones, old rags, that raving slut
Who keeps the till. Now that my ladder's gone,
I must lie down where all the ladders start
In the foul rag and bone shop of the heart.

(Yeats 1939: 236)

7.2.1 The prelude: 'I sought a theme and sought for it in vain'

As it is only theoretically possible to study every single combination in a text, it is up to the researcher to choose which line to compare to its reference corpus counterparts. It is useful, however, to look at those lines which seem to be key transitions in the text, or that otherwise seem to be crucial in the creation of the text's overall impact. In the first stanza, such impact may be created by the first four words as they describe the impasse experienced by the poet: 'I sought a theme'. Below is the concordance of 'sought a' from the Times 1995 reference corpus. The example studies in this section (7.2) is discussed in Simpson (2014: 101-102).

```
MicroConcord search SW: sought a

80 characters per entry
Sort : 1R/SW    unshifted.
    9 ithin the union who have repeatedly sought a compromise with hardliners. One sen
   10 title. Last month Buckingham Palace sought a correction from Business Age over i
   11 ann, for the prosecution, initially sought a court order banning reports of proc
   12  prime radio shows after his mother sought a court ruling proving Julio was his
   13 ty moved to quash that decision and sought a declaration that Mr Oury's applicat
   14 d for the beginning of July. Truman sought a delay until the middle of the month
   15 reated for a nervous breakdown. She sought a divorce on the ground of her husban
   16 ter from her solicitor in which she sought a divorce.   Relatives said Foster su
```

Figure 7.1 The concordance of 'sought a' from the Times corpus.

These are 8 out of the 79 lines yielded *The Times* corpus for the searchline 'sought a'. If we analyse language events in the concordance lines through the Firthian context of situation, we will have noticed a curious consistency: the relevant objects that are being sought, although mostly abstract (legal bans, delarations etc.) like 'theme' in Yeats, are in the vast majority specific in detail. What makes this finding especially interesting is the

presence of the indefinite article in all these contexts – despite or because of which, the object sought is in need of extra description.

In English, therefore, what you seek is very particular and you know what it is very well. All objects of the verb 'sought' in the concordance are specified with the help of relative clauses, prepositional phrases or participle clauses. Yeats's case – or his persona's (they might be seen as one and the same person here, as Yeats' private letters show, see Sarker 2002: 289) – syntactically resembles the situation of the battered woman in line 16 who sought a divorce. (The wider context tells us that she was afterwards killed by her husband on the receipt of her lawyer's letter.) In the whole Times concordance there are eight cases where the sought object is specified by a preceding adjective. There are also three more cases where it remains syntactically isolated. In one, the London Underground sought a solution to a union issue, but the solution is described in a previous paragraph. In another, a married woman, in an attempt to change her life, sought a lover, but in the end went back to her children. There is one more – 'sought a mortgage', the context being a young couple where the boyfriend encounters administrative difficulties in buying their first home because he is self-employed.

A battered wife will seek a divorce, a tired one a lover (this pursuit is less justifiable on moral grounds and perhaps in need of more explanation, but the phrase occurs in a very short synopsis of a short story), and a young couple sought a mortgage. All of these are logical steps, at least no more logical than an elderly poet seeking a theme. The difference is, again, in the specificity of the thing sought. Although no explanation for divorce is given as, for example, in line 15, there is always only one divorce to be obtained; and the very situation of violence in the marriage is sufficient cause for a divorce. A mortgage is also logical when one is buying one's new home, and the particular bank and amount are of no interest in the grand scheme of things. The lover is not described, the very absence of which detail signifies that what was being sought were the usual benefits expected of lovers, personal traits being of secondary importance. But a *theme* – unspecified, and, therefore, posing a problem for the 'relevant person, pesonality' – is a marked deviation from the ordinary usage of 'sought a'. By choosing such a structure to express his problem the artist has foregrounded it at the very beginning (a British National

Corpus search also supported these findings). The semantic prosody of 'sought' is fractured in the poem because the Firthian context of situation is *underprovided*.

The verb 'sought' in the Times concordance above also shows determination on the part of its subject (the determination being highlighted by the very fact that the sought object is usually highly specific). The aura of determination in the poem is preserved by the adverbial 'in vain', which, as Louw (1997) noticed when discussing another poem by Yeats, carries the prosody of a goal-oriented activity.

The determined and unspontaneous nature of the action combined with its highly unspecific object contributes to the creation of the sense of paradox in the very first line. The paradox continues in the second line, where yet again re-enforced 'sought' is specified regarding how long it has lasted ('for six weeks') – the poet is so determined that he keeps a record of the time he has wasted, as if he had an urgent deadline, or was on a payroll and needed to justify his expenses – how often do poets give you this sort of data *within* poetry? After which at the end of the second line, he surprises the reader yet again by adding 'or so', before putting his first full stop. 'Or so' followed by a full stop is not rare even in written English, as the Times corpus confirms by yielding 126 lines containing it, but we would hardly require of Yeats or anyone else such a detailed account of his doings while searching for inspiration. We know that Yeats was far from given to gap-filling for the sake of metre, and yet only a child would finish an iambic verse, and a two-line sentence, with a phrase indicating that the speaker is giving an approximate amount.

It is, of course, quite clear that Yeats' wording is an expression of utter frustration and poetic impasse rather than anything else. What we have found is that the first two lines of the poem are a string of paradoxes which the reader certainly feels, but which a corpus analysis helps to interpret. A determined ('sought' is repeated three times) goal-oriented action whose goal is highly unspecific, because it is unknown, is taking place during a specific time span whose specific duration the poet cannot vouch for.

Let us consider now 'Maybe at last, being but a broken man.' The search string 'but a *ed' yielded five contexts in the Times corpus:

1 Others might argue that the new trend offers conclusive proof that rock culture, as it totters into middle age, has finally succumbed to spontaneous combustion, leaving behind nothing but a charred nylon cardigan.

2 but in winter the theatre is dark, and the pier empty of all but a ragged posse of black-headed gulls, perched on the coin-op telescopes, and a cormorant hanging out its wings on the lifeboat slipway.

3 The rigour of the code, you would have thought, is why a wordless full-page ad showing nothing but a folded shirt lying on a bed of nails turns out to be for Silk Cut. Not so, says the agency behind it. It's been running it since 1983 as a coded message to the mature smoker

4 'Nothing but a naked soul. How frightening!' said the critic Mark Slonim 'The concentration of all female hysterias,' concluded Boris Pasternak. The best critical minds of the day understood Marina Tsvetaeva's talent, but both men and women jibbed at her character. Her

5 But behind the sophisticated visual interface, ISN was more modest.The company was started by programmer Randy Adams and marketing executive Bill Rollinson, with three employees and little to its name but a borrowed Sun server and a link to the Internet.

All these five contexts depict a degraded state of what is expected to be some sort of a more elaborated spectacle. Indeed, Ross (2009: 68) describes the poet's 'accustomed resources' that he can no longer muster as 'the spectacle of his many masks'. Yeats' degraded show is supported by the quasi-propositional variables (QPVs) – 'charred', 'ragged', 'borrowed'. 'Folded' is perhaps an instance of the tricks that the corpus sometimes plays (language is a living thing and not mathematically arranged; a lack of such exceptions would be indicative of an incomplete picture rather than common), but Marina Tsvetaeva's 'naked' soul resonates with Yeats' turning to his own heart in all its unattractiveness (see the last stanza of the poem) so much so that one feels that Yeats is doing as much to explain Contextual Prosodic Theory as it is at pains to interpret Yeats.

The very next line introduces the topic of the poet's heart to which he unwillingly turns ('I must be satisfied with my heart, although…' – the unwillingness is chunked as a state of affairs by the underlined collocates), and the next two lines proceed to inform us that 'Winter and summer till old age began,/ My circus animals were all on show'. What are these circus animals? The sophisticated reader of the sophisticated poet will probably not have thought of real circus animals for the simple reason that he is not reading the confessional of a ring-master. The persona, as we know from the first line, is a poet looking

for inspiration. Therefore, a sophisticated reader will first have understood ‘my circus animals’ as a delexical expression. Then the English expression ‘on show’ will first be understood in its delexical meaning, ‘present’, and only then relexicalised because a show is what happens in a circus. The list of circus animals to follow will only reinforce the process of relexicalisation (see Section 2.5.1). In any case, the ‘circus animals’ in the poem are what they are understood as at first, delexically. They are what a poet, or ring-master, or dancer, or actor must have in order to keep the audience – the gift, the art, the tricks. Sarker quotes Albright (1991: 841) who mentions ‘poetic themes and devices that reliably please an audience, somewhat overfamiliar to the elderly poet who has long employed them’. Apparently, Sarker himself (2002: 290) is closest to the truth when calling them ‘hackneyed images and techniques’. While appreciating this focus on devices (‘images and techniques’) rather than themes, one must not forget that the problem is not their being ‘hackneyed’ or ‘overfamiliar’. The problem is that old themes do not work any more because Yeats’ animals are gone. As critics agree that the reference to lion and woman is reminiscent of Yeats’ unrequited love Maud Gonne, I am correct in assuming that she, as the source of inspiration, is gone too.

7.2.2 The middle stanzas: ‘What can I but enumerate old themes’

What can Contextual Prosodic Theory contribute to existing scholarship on the three middle stanzas? It can show the subtext of the middle section of the poem and talk about devices connecting the studied text and the reference corpus. Now that literary devices have deserted the poet, he feels driven into a corner. He sees the scene of his utter plight and helplessness, as the empty stage is refilled by personas of the past. This is not an intuitive statement; the words in italics in the concordance below represent scientific findings based on the extracted subtext of the grammatical string with which Yeats begins the second section of the poem: ‘what can I but *’.

The grammatical string in question is foregrounded by the mere virtue of its being extremely rare – archaic. The *Times* corpus, the British National Corpus (twice its size), and COCA do not contain any examples of this line (all searches I mention in this section

were made in September 2012). We will see that the use of such grammatical strings is one of this poem's characteristics.

Only the Google Books corpus (googlebooks.byu.edu) assisted the investigation. The Fiction part of the corpus containing 91 billion words provided only seventeen – as far as I could find, as many passages are repeated on the Google Books findings pages – contexts containing 'what can I but *'. The QPVs in these contexts are 'despair' (3 occurrences), 'obey' (2 occurrences), 'repeat' (2 occurrences), 'love', 'plead', 'pray', 'say', 'see', 'submit', and 'tremble'. Four contexts are translations, one into modern English. Here follows the KWIC concordance of the contexts (see Figure 12.1 in Section 12.5.2 for opened contexts of this concordance):

1 those who won't receive it – *what can I, but* ***say*** my say too, & trust in God if I'm wrong
2 He flops, cries 'stand:' *what can I, but* ***submit***? To fight a drunken bear, were want of wit
3 What can I do – let doubt be dumb, - *What can I – but* ***obey***? His sceptre or His rod, Who sh
4 full and just Inspires my words – *what can I but* ***obey***? Yet as I stand, held fast and meshe
5 Look there On my young wife! *What can I but* ***despair***? She left her tents for me — abandon'd
6 thus oppressed with endless pains, *What can I but* ***despair***?' Then rising, hopeless of relie
7 If all my thought me helpeth nought, *what can I but* ***despair***? Sorrow and sighs and dreary m
8 all the strength of my heart, and *what can I but* ***pray*** for thy whole blessedness, O gentle
9 But if Destiny to this consent, *What can I but* ***my soul in sorrow bow***, With tearful eyes an
10 to your imagined throne? And *what can I, but* ***see*** beyond the world that is, when, faithful
11 It is not, *What can I? but*, ***What can't He?*** as somebody says. Go on fearle
12 nor like the flute complain? *What can I but*, ***like the ended banquet, desolate remain***? 'Fe
13 sigh for you; And knowing this, ah! *What can I But* ***love*** in silence, pine and die?
14 outcast from natural pride, *what can I but* ***plead*** the greater love I bear you as my benefa
15 *Nina* Do not cease! Say on, say on! *Doria* *What can I but* ***repeat*** A tale already told? *Nina*
16 far silence touched it. *What can I but* ***repeat*** The vow of every mother – There is not one
17 this swoop makes the third — And *what can I, but* ***tremble*** like a bird? FOOL. Give me a pen

Figure 7.2 KWIC concordance of the contexts of 'what can I but *' from the Google Books corpus (googlebooks.byu.edu)

The QPV 'despair' appears three times. In context 5 the husband is desperate as his young wife is dying and leaving their first new-born; in 6 a maiden has suffered at the hands of a man whom she still loves, and in context 7 a man is suffering from what we now call 'courtly love' (this is a translation into modern English).

'Obey' appears in 3 in a hymn celebrating the omnipotence of God, and in 4 where a subject is addressing his Queen.

The QPV 'repeat' beautifully resonates with 'enumerate' in meaning, with two differences: one being that 'enumerate' has a routine connotation, and the other that Yeats

uses it to state a repetition that is not welcome. ‘Repeat’ is used in the context of man-woman love in context 15, and motherly love in context 16.

‘Love’ comes up in a context of unrequited love with the traditional elements of the inaccessibility of the beautiful beloved; ‘plead’ has to do with a princess pleading with her sovereign, who is in love with her, not to marry her; ‘pray’ is self-explanatory; ‘say’ is from a letter from Thackeray to his mother, in which he asserts his faith - different from hers. ‘See’ has to do with a mother reading fairy tales to her daughter and seeing the world beyond reality; and ‘tremble’ comes from a play by none other than W.B. Yeats himself and as such will deserve special attention.

There are also two contexts (9 and 12) in which the verb appears further to the end of the clause (both are translations). In another context (11) there is a question mark after ‘I’ and a comma after ‘but’, but we assumed it has a right to its place in the concordance because of Firth’s notion of collocation as abstracted at the level of syntax. These contexts will be consulted when the prosody of the string is discussed, but not checked for QPVs.

According to Louw (personal communication), the QPVs will not necessarily have one semantic feature in common, but rather each two of them will share one. The QPVs from this grammatical string can be grouped semantically in the following way:

1. ‘despair’
2. ‘obey’, ‘plead’, ‘pray’, ‘submit’, ‘tremble’
3. ‘love’
4. ‘pray’, ‘plead’, ‘say’, ‘repeat’

They can also be grouped thematically according to their specific contexts:

1. despairing of love: ‘despair’ (all three times), ‘love’
2. asserting faith in God: ‘pray’, ‘obey’ (once), say
3. asserting love: ‘repeat’ (both times)
4. new vision: ‘see’, ‘tremble’
5. humorous: ‘submit’

All contexts can be grouped into three major themes: despair (5, 6, 7, 9, 12, 13, 17), faith in God (1, 3, 11, 17), love (5, 6, 7, 9, 13, 14, 15, 16), and submission (3, 4, 9 (to fate), 14, 17). New vision is strongly present in 10 and 17.

The play *The Hourglass*, where Yeats himself uses the grammatical string in question, deserves particular attention. It is about a transition from atheism to faith. A schoolmaster (in the play called Wise Man) has asserted all his life that there is no God, nor angels, spirits, hell etc, condemning it all as superstition. He is widely respected and no one contradicts his views, including his pupils, wife and fellow villagers. However, he begins having dreams and visions that he has been wrong – that is the moment when this realisation prompts him to 'tremble'. God exists, and suddenly Wise Man is the greatest sinner through non-believing and teaching it to others. It is obvious that the usage of the grammatical string in *The Hourglass* expresses despair, faith in God, new vision and submission. Love is not at all mentioned in the play, and therefore, cannot be considered as subtext of this line in *The Hourglass*.

Let us, however, put into words the corpus-based description of the subtext of 'what can I but *' in general terms. Provided the context of situation contains elements that warrant these headings, this grammatical string will conjure up contexts of love, despair, faith in God, submission and new vision. 'Love' is likely to be unrequited or lost (even in context 15 Nina tragically dies, unable to bear the hero's adultery). Submission is likelier to be directed to God. Despair is likelier to be caused by love and not, for example, fate, as in context 12. Interestingly, the subtext of this string bears resemblance to 'faith, hope, and charity' (St Paul: 1 Corinthians 13), with the correction that charity is likely to be love unfulfilled, and, not illogically, the place of hope has been taken by its opposite, despair. When I entered 'what can I but hope' into the Google Books corpus, there were no matches.

Far from claiming that all these will be consciously used by the writer or clearly perceived by the reader, Contextual Prosodic Theory nevertheless asserts that subtext deepens interpretation through being embedded in the grammatical string in the nature of the language experience of an educated adult native speaker. The theory is called

'contextual' because it relies heavily on contexts and *events* in both the studied text and reference corpora (for example, 'love' is not the subtext in *The Hourglass* because the context does not warrant it). It is called 'prosodic' because it relies on collocation, within which words imbue each other with meaning, as in the case of semantic prosodies, but also within the limits of a one-time co-occurrence. Thus, the theory is both dynamic and precise.

Let us now view the subtext of Yeats' line in 'What can I but enumerate old themes'. The archaic grammatical string collocates well with 'old themes'. According to Louw (2010b), there will be a grammatical string in the poem that will prove to carry the poem's subtext. Does the string I have been researching carry the subtext of the poem, or at least of its middle section?

Considering the context of the middle section, the subtext carried by this grammatical string is that of love (middle stanza 1), faith (middle stanza 2) and despair (middle stanza 3). Clearly the elements of subtext are not restricted to each stanza, as the second stanza is also about love, and the first about despair (it contains our string). Is love in the string's subtext unrequited? Clearly so. Has this fact produced despair in the form of writer's block? Not unlikely. Is the presence of faith in the subtext strong enough to be taken into consideration? The answer to this remarkable question is present both in stanza 1 and in stanza 3.

Stanza 1 famously ends with the delexical expression 'the Lord knows what'. Yeats never wasted words; the poem is multilayered; this delexical expression ends the whole prelude. If we accept the possibility that this expression is not used delexically only, there will be an element that will relexicalise it. The element in question could be the very string I have just analysed. The poem, in fact, runs as follows: '…and the Lord knows what. What can I but …' In the briefest possible way, 'the Lord' is relexicalised through the string that even Yeats himself has used in the past before the QPV 'tremble'.

7.2.3 The coda

The concluding stanza presents a change of focus. If in the preceding stanzas the focus was on what has departed, here the focus is on what has been left. If the prelude gives a mocking list of devices, and the middle stanzas a respectable description of lofty old

themes, the coda suddenly shifts focus from ‘the pure mind’ to the squalor of the poet’s heart:

> Those masterful images because complete
> Grew in pure mind, but out of what began?
> A mound of refuse or the sweepings of a street,
> Old kettles, old bottles, and a broken can,
> Old iron, old bones, old rags, that raving slut
> Who keeps the till. Now that my ladder's gone,
> I must lie down where all the ladders start
> In the foul rag and bone shop of the heart.

Similarly to the mocking list of animals, the coda contains the exhaustive to the point of over-provision description of utter waste that clutters the poet’s heart. Unlike tricks, whether they be circus or poetic, no item of this detritus could ever serve a reasonable purpose to anyone but the rag and bone man (overprovision of uselessness) or attract so much as even a condescending smile (overprovision of disgust). Yeats’ public would welcome tricks but not disgusting junk. All items in the shop are invested with the attribute of being old, like the poet himself, except ‘the raving slut’ (the ‘broken can’ is not described as old but this is easily inferred as it is broken). Interestingly, both previous sections ended with an important message whose last word and carrier of the rhyme was a grammatical word (‘the Lord knows what’, ‘emblems of’). In the case of the concluding stanza, the last word of the concluding message and the carrier of the rhyme is the word ‘heart’. The literary world of the last stanza is, however, more surreal than in the prelude: if there could exist a circus corresponding to Yeats’ description, there could hardly exist a junk shop, however poor, that was run by a crazy person into the bargain (‘raving’ and ‘keeps the till’ are impossible collocates).

The progress from a reasonable and even cool comparison of literary devices to circus animals, via a sensible account of one’s artistic development (containing sound judgment like ‘counter-truth’ and ‘heart-mysteries’) towards a surreal mess of the poet’s heart, which yet needs to rhyme with ‘start’, is paradoxical. This reminds us of the

numerous paradoxes of the first stanza; however, the paradoxes are as opposite as the stanzas' moods. At the beginning the poet is so calm and collected as to insist on giving us an exact estimate of time lost on writer's block; the list of animals that follows completes the process of relexicalisation because its completeness recreates the world of the circus. By comparison, the list of items belonging to the junk shop (ostensibly having the same purpose of rounding up a metaphor) is not quite coherent. 'A mound of refuse and the sweepings of a street' are an exhaustive and disgusting enough description, but are they part of the junk shop? Hardly; rather, they are the summary of it, an exaggeration preceding the actual description of what is in itself a metaphor. Then, Yeats (the persona) must lie down in the shop ('shop' and 'lie down' are unlikely collocates). Before that, the appearance of a ladder in any shop is very likely if its point is to reach a shelf, but it is never its central feature. Ladders usually do not start (unless metaphorically); this one should obviously lead to heaven as there is nothing else in that direction.

This picture of utter underprovision (Louw 2000) is yet another 'masterful' image, 'because complete'. The image of the shop is so physical due to overprovision of disgusting items that the appearance of a 'ladder' – a traditional metaphor of spiritual growth – is almost palpable, as well as the author's intention to 'lie down' in the shop. If in stanza one the animals were first delexically present, and then relexicalised, here the literary world of the shop relexicalises the metaphoric notion of the ladder by co-selecting it with the till and the shop junk that have become a world. The 'ladder' is an absent collocate that becomes equated with the missing animals and the focus of the poem – in fact, its theme. The equation comes from the statement that they are both gone. Ladders do not start and not in junkshops, except metaphorically and of the heart, but by the time we have reached the mention of this one, we have taken to heart the loss and the point of the poet' new departure because of the 'completeness' of his created worlds. If the poem started with the 'ladder' metaphor, without preliminaries, the impact would have been poorer as the image in itself is almost conventional.

Let us now turn to the subtext of the stanza. Given the findings of the previously analysed string, the subtext of which was hypothesised to be love (unrequited or lost), despair, faith, new vision and submission, I will ask two important questions. Firstly, will

the subtext of the middle stanzas prove plausible? ‘Lie down’ certainly implies submission, and ‘all ladders start’ certainly implies faith, as well as new vision (it will not be amiss at this point to recall St John of the Ladder). We are left with the task of proving or doubting the presence of lost love and despair. The second question stems from the first. Is the raving slut of the coda Maud Gonne?

Is the image of the slut Yeats’ judgement on Maud? The poem is multi-layered, there are numerous parallelisms, and this hypothesis is not ruled out. A stylistician, however, will notice two points. Firstly, ‘my dear’ in the third stanza refers to Maud and the transition from ‘my dear’ to ‘slut’ would be too unjustifiably fast, even if we take into account the ‘flashback’ aspect of ‘my dear’. Secondly, ‘the raving slut’ is the first item of the workshop that does not come with the epithet ‘old’. In such a repetitive list describing the old poet’s heart nothing would contradict the appearance of ‘old’ instead of ‘that’ except for (a) the need for the feminine image to be truly feminine, not after child-bearing age, and (b) the possibility that the reading public might *mistakenly* read Maud Gonne into the image. She was past the age of 70 at the time.

The subtext of the coda may be determined on the basis of two search lines. The grammatical string that has a bearing on the notion of the junk shop is ‘but out of what’. The grammatical string that will elucidate the degree of the poet’s despair and the presence (or absence) of the motive of lost/unrequited love is ‘now that my * is’.

It is Louw’s view that enough empiricism (namely, similar events in the corpus) allows ‘text to read text’ and that *the start of the event can* predict *its progress or* subvert *it into device* (Louw and Milojkovic 2016: 201; see also Louw 2008 on literary devices). The third possibility is insincerity (Louw 1993, see also Section 5). The findings will show that the former grammatical string (‘but out of what’) is an example of the event being *predicted* by corpus findings. The latter (‘now that my * is’) will be shown to *subvert* the event into a literary device – metaphor.

‘But out of what’ is a rare grammatical string. As in the case of ‘what can I but’, neither the BNC nor COCA contain any examples of it. Only Google Books: Fiction, a corpus of 91 billion words, yields material for study. Fifty-four contexts were extracted from it.

Since ‘but out of what’ is syntactically capable of being followed by both verbs and clauses, the range of the potential QPVs is so huge that we should concentrate on its semantic prosody (aura) rather than subtext that is extracted through the variables of lexis. The contexts of ‘but out of what’ were grouped into 7 categories, according to the semantic context of their occurrence. The categories are given in order of frequency, starting with the most frequent. Since three contexts quote the same thought of the same writer, I will consider the overall number of contexts to be 52:

(a) creation (15/52, or 28,8%):

- artistic creation in 5 (the same in 13 and 46), 8, 17, 29, 51, usually followed by
    - underprovision of resources (‘out of what you don’t know’ (5, 13, 46),‘out of what they most lacked’(17)) or
    - underprovision of the knowledge of the resources’ nature (the rhetorical question ‘but out of what?’ (6)), ‘out of what people are unpredictably going to say or do’ (51)
- Christian contexts in 7, 30, 31, 39, 54
- philosophy in 1, 6, 26; anthropology in 6 and architecture in 43.

(b) financial transactions, as in taxes and other payments (11/52, or 21.2%) in 2, 10, 11, 12, 15, 23, 47, 48, 50, 52, 53; underprovision of resources is implied in 12, 23, 47, 48, 52 (5 contexts out of 11); in 2, 10 and 11 the theme is charity.

(c) government and public affairs in 3, 9, 14, 27, 37, 42, 44 (7/52, or 13.5%).

(d) hedging in 18, 19, 32, 33, 34, 35 (6/52, or 11.5%), e.g. ‘out of what is called’, ‘out of what is best described as’, etc.

(e) unknown origins in 20, 21, 25, 36, 38, 45 (11.5%).

(f) rhetorical questions implying the absence of resources (the extreme form of underprovision) in 16, 24, 28, 41 (4/52 or 7.7%).

(g) origins of socially undesirable behaviour in 22 and 40 (2/52 or 3.85%).

Two types of conclusion are to be drawn from these findings. One is that grammatical strings show specific prosodies, again confirming the empiricism afforded by corpus-derived subtext; therefore, in a text they may denote the beginning of a *predictable event*.

The other conclusion concerns Yeats' art. It is remarkable that he was a poet of such depth and awareness of his own language as to unwittingly (intuitively) employ a grammatical string whose main prosodies are creation and (underprovided) payment, at a point when he was introducing a metaphor of origins of creation presented as a junk shop. Paradoxically, the subtext (or semantic aura) of the string emphasises both optimism (creation) and pessimism (insufficient or unknown resources).

The other chosen grammatical string, 'now that my * is' turned out to subvert the event into a literary device – metaphor. There are only two such contexts in the BNC. One is religious: 'My strength is waned now that my *need* is most'; the other comes from fiction: '… but now that my *grandmother* is feeling better…' One is underprovided, the other more positive. COCA has yielded 17 concordances (those examples where 'now that my * is' was followed by a grammatical word, e.g. a preposition, were considered different grammatical strings and were not included):

1 Now that my father is dead
2 Now that my father is dead
3 now that my father is far away
4 now that my job is done
5 now that my job is our sole source of income
6 Now that my hair is growing back
7 Now that my hair is turning gray
8 now that my youngest is started school
9 now that my wife is aware
10 Now that my sister is single
11 now that my Sandy is gone.
12 Now that my office is 10 miles from where I live
13 now that my mind is glutted
14 Now that my inability is real
15 Now that my house is uninhabitable
16 Now that my brain is empty
17 now that my beloved is gone?

Figure 7.3 COCA concordance for 'now that my * is *'

'Father' appears three times, 'hair' and 'job' appear twice. The quasi-propositional variables may be classified into the following subgroups:

(a) a family member or a very significant person in 1, 2, 3, 8, 9, 10, 11, 17
(b) body part in 6, 7, 16
(c) significant quality in 13, 14
(d) task in 4 (‘job’) and 12 (‘office’)
(e) significant basic necessity in 14 (‘house’ co-selected with ‘uninhabitable’ makes the QPV dramatically necessary; ‘office’ could also be mentioned here).

Indeed, all the QPVs are those of basic significance to the person (personality in Firth’s (1957: 182) context of situation). If we turn to the corpus of Google Books: Fiction for more data, these are the QPVs yielded: ‘father’ (499 occurrences), ‘mind’ (186), ‘life’ (167), ‘mother’ (146), ‘heart’ (126), ‘son’ (117), ‘time’ (110), ‘husband’ (107), ‘brother’ (104), ‘hair’ (87), ‘innocence’ (84), ‘wife’ (73), ‘turn’ (64), ‘task’ (63), ‘name’ (56), ‘hand’ (55), ‘work’ (54), ‘daughter’ (51), ‘master’ (51), ‘head’ (47), ‘temper’ (46), ‘honour’ (44), ‘presence’ (41). The classification, if attempted, would correspond to the one suggested in the discussion of COCA findings.

It would be statistically more valid but too time-consuming to check which of these occurrences are followed by a grammar word (thus forming a different grammar string) so that they could be excluded. At the time of writing it was not possible to suggest to the corpus a search string longer than 5 words. When I attempted to add ‘gone’ to it, for reasons similar to those of including ‘made’ in Louw (2010a), I needed to do it manually. This means I could only ascertain which QPVs become invalid after the addition of ‘gone’. The invalid ones are: ‘work’, ‘task’, ‘turn’; ‘head’, ‘hand’; ‘temper’, ‘presence’. This leaves (and it is not clear in which order of frequency, therefore I have preserved the one yielded by the first search):

father mind life mother heart son time husband brother hair innocence wife name daughter master honour

These QPVs share one quality: extreme meaningfulness to the owner. If classified, there are two sub-groups:

- a very significant person

- an essential quality without which life or honourable existence is impossible.

Who or what is the ‘ladder’? It is Maud Gonne? And gone she is indeed, leaving the raving slut in her place and causing greater despair than is ostensibly stated in the text. The relevant person or personality in Yeats’ art whose effect was his verbal action is felt, directly or indirectly, in the first four stanzas of his poem, before appearing as subtext in the coda. She appears as a double absence, by not turning up in the line directly, as something like ‘Now that my beloved is gone’, or less directly – ‘Now that my inspiration is gone’, and by being co-selected with the past participle ‘gone’, homonymous with her last name. Yeats is known to play with Maud being Gon(n)e; to what degree the pun is a part of his conscious authorial intention, we cannot tell – nor is it essential in the presence of subtext. He is about to create in a Maud-free world, despair truly great, love truly lost. No wonder he ‘must’ ‘lie down’. There is evidence, however, that the circus animals, in a different form or shape, have announced their return. The evidence in question is the present poem.

7.3 Subtext in the key metaphor of 'The Circus Animals Desertion'

The previous section has shown how corpus-derived subtext can either confirm the meaning that is intuitively deducible from the poem (e.g. 'but a *', 'but out of what'), or further inform the reader's understanding of it ('sought a', 'what can I but *', 'now that my * is gone'). However, not much proof has been offered of the interrelatedness of semantic auras within the poem, except at a general level that the reader can identify when processing the poem's message. This section will offer proof that interrelatedness of grammar strings within a text is an empirical phenomenon. Since the grammar strings in question will both refer to the central metaphor of the poem, this section will end in an attempt to offer a corpus-attested definition of metaphor as a rhetorical device, based on the two instances discussed.

Clearly, the main metaphor of ‘My Circus Animals’ Desertion’ is inherent in line 6 of the first stanza:

> Winter and summer till old age began
> **My circus animals were all on show**,

Those stilted boys, that burnished chariot,
Lion and woman and the Lord knows what.

The meaning of ‘on show’ is undoubtedly delexical. ‘Circus’ and ‘show’ are separated by four words, which is in accordance with Sinclair’s nine-word window of collocative power (four words to the left of the node and four words to the right). This means that ‘circus’ and ‘show’ exercise collocative power over one another. If we co-select these two words in the British National Corpus (BNC), under the same conditions (4 words at most between them), we will find that they all the state of affairs of a circus show (the reader is referred to COCA).

As previously stated, Contextual Prosodic Theory treats the reference corpus as a sample of the world. Simultaneously, and closer to home, the reference corpus is taken by it to represent the language norm. Since the reference corpus is balanced and representative, and thus reflects the language it is a sample of, we may assume that it more or less represents the accumulate language experience of its speakers. This is confirmed by tendencies in language use that are consistent but cannot be intuitively accessed, such as SP and subtext. Therefore, if ‘show’ relexicalises in the presence of ‘circus’ in the reference corpus, it must relexicalise for the native speaker. Also, intuitively we conclude that ‘circus animals’ must mean ‘literary devices’, or something to that effect: they are the ‘topic’, and ‘circus animals’ are the ‘vehicle’ (this terminology is used in accordance with Philip 2011: 75). If we look for a corpus-attested confirmation of this hunch, we can co-select ‘device*’ with ‘theme’ (‘theme’ being the issue under discussion in the stanza) and find the following single context in the BNC: ‘Eikhenboum’s essay on Gogol’s *Overcoat* is a good example of the latter. He sets out to show that in Gogol’s tale the centre of gravity is transferred from the **theme**… to the **devices** (1963: 377)’. COCA yielded five contexts, in which those of artistic creation predominated:

1 what I call the ‘shaggy-dog story’ **device**. The opening **theme** of the Eroica be
2 ariation on the solar **theme** is the **device** developed in Florida Solar Energy c
3 family (of TV and movie fame) as a **theme**. The complex **device** ended when Thing
4 Wood’s use of the seasonal **theme** as a **device** for expressing the doctrines of
5 Besides these overt lexical **devices**, the **theme** on which the plot hinges, enco

Figure 7.4 'Theme' and 'devices' co-selected in COCA

Co-selection chunks the states of affairs of composing music (line 1), technical inventions (lines 2 and 3), painting (line 4) and lexical devices in American SF (line 5). Thus, it is possible to establish the implied subject of the metaphor through co-selection and the states of affairs in the corpus.

The contribution of this section to (corpus) stylistics is the tentative assumption that corpus-derived subtext generally seems to play a major role in Yeats's metaphors, albeit invisibly. The descending frequency list of the items in the lexical slot of the line 'my * were all' in the Google Books-UK corpus is found below:

1 my **friends** were all (156)
2 my **thoughts** were all (136)
3 my **children** were all (64)
4 my **companions** were all (60)
5 my **eyes** were all (59)
6 my **nerves** were all (56)
7 my **men** were all (53)
8 my **clothes** were all (49)
9 my **sympathies** were all (46)
10 my **family** were all (45)
11 my **hands** were all (41)
12 my **dreams** were all (40)

Figure 7.5 The concordance of 'my * were all' from the Google Books-UK corpus

The QPVs of 'my * were all' can be divided into four categories:

I	II	III	IV
friends (156)			
	thoughts (136)		
children (64)			
companions (60)			
		eyes (59)	
	nerves (56)		
men (53)			
			clothes (49)
	sympathies (46)		
family (45)			
		hands (41)	
	dreams (40)		

Judging by the wider contexts from the reference corpus, the first category can be described as ‘significant people in one’s life, such as family and friends’. The second is ‘thoughts and emotions’ (with one exception as far as I could see, that is the meaning of ‘nerves’ in the reference corpus). The third, 'eyes and hands' is significant because these organs we need to see and do things. The fourth, ‘clothes’, is in the reference corpus often in bad condition, and may endanger the health of the narrator. While all the four categories are significant, the first two have a direct bearing on the poem.

Let us start with the second category. In order to illustrate the presence in the poem of this subtext, it is enough to read through the subsequent stanzas (the context clues related to thought and emotion are highlighted n bold):

II
What can I but enumerate old themes,
First that sea-rider Oisin led by the nose
Through three enchanted islands, allegorical dreams,
Vain gaiety, vain battle, vain repose,
Themes of **the embittered heart**, or so it seems,
That might adorn old songs or courtly shows;
But **what cared I** that set him on to ride,
I, **starved for the bosom of** his fairy bride.

And then a counter-truth filled out its play,
`The Countess Cathleen' was the name I gave it,
She, pity-crazed, had given her soul away
But masterful Heaven had intervened to save it.
I **thought** my dear must her own soul destroy
So did fanaticism and hate enslave it,
And this brought forth a **dream** and soon enough
This **dream** itself had all my **thought** and **love**.

And when the Fool and Blind Man stole the bread
Cuchulain fought the ungovernable sea;
Heart mysteries there, and yet when all is said
It was the **dream** itself **enchanted me**:
Character isolated by a deed
To **engross the present and dominate memory**.
Players and painted stage took all my **love**
And not those things that they were emblems of.

III
Those masterful **images** because complete
Grew **in pure mind** but out of what began?
A mound of refuse or the sweepings of a street,
Old kettles, old bottles, and a broken can,
Old iron, old bones, old rags, that raving slut
Who keeps the till. Now that my ladder's gone
I must lie down where all the ladders start
In the foul rag and bone shop of the **heart**.

The context clues, highlighted in bold, are convincing proof that subtext does relate to the aspects of the text that are accessible to our intuition, and that it is not random or separate from the grammar string's substantiation in the language. In fact, all the four categories of the subtextual lexis point to the utmost significance of what has 'deserted' the elderly poet. And that is where we would leave it, had it not be for very substantial evidence that there is more justification to the first category ('significant people in one's life') than mere significance in general terms.

Let us look at the underlined text in the coda, discussed at length in the previous section: ‘Now that my ladder's gone... ’ This is a metaphor; moreover, ‘ladder’ obviously implies the same entity as ‘circus animals’ in stanza 1, because of the collocate ‘gone’. Section 7.2.3 provides an extensive corpus analysis of the grammar string ‘now that my * is’ in combination with ‘gone’. The subtext reveals two sets of lexical variables: a) a

significant person (‘father’, ‘mother’, ‘son’, ‘husband’, ‘brother’, ‘wife’, ‘daughter’, ‘master’, and b) a quality essential to (honourable) existence (‘mind’, ‘life’, ‘heart’, ‘time’, ‘hair’, ‘innocence’, ‘name’, ‘honour’). The ‘ladder’ is also relexicalised in the context of so many tangible objects located in the rag-and-bone shop. Clearly, two metaphors in two very different sections of the poem (the opening stanza and the coda), employing different grammar strings when referring to the same thing, contain the same or very similar subtext. Based on these two examples, we can give a working definition of Yeats's metaphor:[36]

> Yeats's metaphor consists of a grammatical string containing a lexical item unique for the string in question (not found in the reference corpus).
>
> The lexical item will interact with the surrounding context clues in two ways:
>
> - relexicalisation will allow it to attain its literal meaning in the text and the reference corpus
> - co-selection will allow it to attain a second, implied meaning in the text, accessible at the intuitive level, which is the purpose of the metaphor; the states of affairs chunked in the reference corpus by co-selection will confirm this
>
> The most frequent lexical items of the grammatical string in question (its subtext) will represent a third, hidden meaning. They will interact with the context clues in ways which will support the existence of the second, implied, meaning.

The two metaphors studied in this section have something in common, besides the same (approximately) implied meaning and very similar subtext. The ‘vehicle’ is, in both cases, a relexicalised entity of a much lower status than the implied. Circus entertains on a much more primitive level than poetry. The ladder, despite the intertextuality (St John of the Ladder), is too sudden and, because of its collocates, too physical to be viewed respectfully – together with the poet who, we are told, used to be in need of one, and now has to do

[36] This definition takes into account original metaphors, as in Yeats, unlike delexical metaphors discussed in Lakoff and Johnson (2003).

without. This depreciation, especially in the context of the title (the metaphor in the title is that of the circus animals, and not that of the ladder) is certainly tinged with self-irony – but given the extent of the defeat, the humour is bitter even if ‘ladders’ are co-selected with ‘start’ towards the end.

7.4 Pun

So far it has been shown that both inter-related metaphors in 'My Circus Animals' Desertion' contain wordplay: 'circus animals' co-selected with 'on show' in its delexical meaning, and relexicalised 'ladder' co-selected with 'gone'. It would be useful at this point to attempt a corpus-attested definition of pun to see how it differs from that of (Yeats's) metaphor.

These two definitions of pun can be found in traditional sources: ‘a figure of speech which involves a play upon words’ (Cuddon and Habib 2013:572), and ‘a play on words that are either identical in sound (homonyms) or very similar in sound, but are sharply diverse in significance; an example is the last word in the title of Oscar Wilde’s comedy, *The Importance of Being Earnest*’ (Abrams and Harpham 2009: 295). Despite the difference between these definitions, both sources give the following two examples, among others: John Donne, ‘Hymn to God the Father’ (‘And having done that, Thou hast done’), and William Shakespeare, *Romeo and Juliet* (Mercutio: ‘Ask for me tomorrow, and you shall find me a grave man’). In the first case, the pun is sustained through the homophony of ‘done’ in the meaning of ‘finished’ and ‘Donne’ as in John Donne: God will accept Donne’s soul after he has forgiven all his sins. In the second example, the dying Mercutio plays with the meaning of ‘grave’ as a noun and ‘grave’ as an adjective, while the co-text of the pun contains discussion of his mortal wound.

The following example of wordplay is from a novel by Charles Dickens (1993: 368):

> ‘I say, old boy, **where do you hang out**?’
>
> Mr Pickwick replied that he was at present **suspended** at the George and Vulture.

In this context, ‘where do you hang out’ is a fixed expression, meaning ‘Where are you to be found?' The word ‘suspended’ is never used in this context, but it does reactivate the literal meaning of ‘hang out’, as in ‘to hang out the washing’ (Cobuild 1998).

An example of pun in modern poetry can be found in Philip Larkin's ‘Winter Nocturne’ he describes the advent of night as unrest in nature. The poem finishes with

The rain falls still: bowing, the woods bemoan;
Dark night creeps in, and **leaves the world alone**.

The whole poem is about the period of transition between day and night, culminating in its arrival. How then can it at the very moment of arrival leave the world alone? Apart from the meaning of ‘leave something in a particular state’ (which implies that the world is alone now night has entered it), Larkin must also have had in mind the meaning ‘not bother or interfere with, leave be’. The word ‘leave’ co-selected with ‘alone’ was searched in the 1995 Times newspaper corpus. A hundred and fifty concordance lines were found. In fifteen of them (10%) the context was literal (e.g. leave a child alone, i.e. without adult supervision). Out of the remaining 135 delexical usages, 99 (73%) pointed to the conclusion that leaving something or someone alone in the sense of not interfering with it is a good idea, because of the unwelcome consequences of not leaving alone. Therefore the poem about the advent of night describing unrest felt in nature prior to its arrival states clearly that night, by leaving the world in the state of loneliness, does a good thing (Milojkovic 2011a).

Puns are often the basis of comic one-liners, e.g. 'Santa’s helpers are called **subordinate clauses’.** Here ‘subordinate clauses’, a frequent collocation in the reference corpora, becomes viewed as a combination of ‘subordinate’ in the sense of ‘less important’ (Cobuild 1998) and ‘clauses’ in the sense of ‘little santas’. Puns also often feature in advertisements, or literary titles – for example, *Brothers in Law* is a comic BBC serial about an idealistic young lawyer (http://www.bbc.co.uk/programmes/b050gvnz, accessed on 29 November 2015).

If we look at these examples from different genres, it will be clear that the majority contain a frequent collocation, one of whose constituents undergoes relexicalisation. Therefore, the following corpus-attested definition emerges:

> Pun is characterised by a double meaning of a lexical collocate which is part of a frequent and often delexical collocation, whether lexical ('grave man') or lexico-grammatical ('thou hast done'). The double meaning is composed of the meaning of the lexical item as part of the frequent and delexical collocation, as found in the reference corpora, and a provisional meaning, often literal, imposed on it by context clues (the effect of relexicalisation), whether immediate ('Santa's helpers are known as subordinate clauses'), distant (the discussion of Mercutio's wound) or extra-textual (the surname Donne is not in the poem where the pun appears).

This definition should be checked, and probably refined, against other existing examples of pun. At this juncture, however, it has already become obvious that collocation, as understood by Sinclair (as co-occurrence within a certain distance in the text, the distance being measured by words) does not wholly determine this literary device, at least not on all occasions. 'Santa's helpers are called subordinate clauses' is a perfect example of Sinclair's nine-word window, with exactly 4 words intervening between 'Santa's' and 'clauses'. But that does not have to be the case: in some cases it is necessary to have the knowledge of the macro-context, and even external factors (the author's name, for example) in order to appreciate to the full the effect of pun.

This definition has been arrived at in the top-down way, in the sense that it seems to work in cases of what I *know* to be wordplay. It is to be hoped that it also works bottom up, and an independent entity not possessing intuition, such as the computer, could highlight the cases of pun according to the definition's instructions. But what is the actual source of humour? When is the computer to laugh (let alone smile)? It appears that, since the provisional, literal meaning is the new element in the arrangement, whereas the delexical meaning is the one normally used in the language and found in the corpus, it is the provisional meaning that provokes laughter.

Some jokes, however, are only funny if you are familiar with their social or political context, or if you hold certain political beliefs, and this is when Malinowskian context of culture comes to mind. But solely on the basis of these examples, we can distinguish between serious ones (Donne, Larkin) and humorous one (Dickens, the BBC series title). It appears that the impression of incongruity prompts laughter, while if the collocation simply fits the context in two different ways, the reader is appreciative rather than amused. At the same time, the point of a one-liner does not go beyond the laughter-inducing quip. 'Subordinate clauses' are their own point, there is no broader context where they play a part if literally taken. Judging by Donne and Larkin, double meaning in which both senses contribute to the message seems to emerge in cases where laughter is expressly out of place. But even in the latter cases, one feels that if the author has risen to wordplay, the tragedy is not absolute.

The computer, then, may be taught not to laugh if it registers certain context clues that preclude laughter. The point of this lengthy discussion, some of whose premises have already been stated elsewhere, is that the scope of figures of speech may go far beyond their immediate contexts.

Subtext, being opaque to intuition, cannot have humorous implications, unless very subliminally. In fact, it may have a completely different semantic aura and still not interfere with the jocular implication. There is no mention of subtext in the offered definition of pun. There is a reason for this: although examples of wordplay, even if one-liners, must contain lexico-grammatical collocatons and therefore subtext, the subtext of the grammar strings in question can never be a rightful participant in creating puns, provided they are their own goal. An example of subtext in puns is discussed in Louw and Milojkovic (2016). The joke in question is Hot Tomato:

Q: Why did the Tomato blush?

A: Because he saw the salad dressing.

As it turns out, the grammar string 'he * the * *ing' contains a frequent lexical item, or quasi-propositional variable, 'saw'. However, the semantic aura of the string clashes with the humour genre, because it appears either in religious contexts, or in those describing a

(dangerous) turning point (Louw and Milojkovic 2016: 138-140). It may be that such contexts of situation always inspire more consideration than, on the face of it, is afforded them by their observers or even participants (ibid.). Our concern at present is to state that, although subtext is always present in a lexico-grammatical collocation, it cannot have a point in 'word-play', as the very term suggests. Puns serve a different purpose altogether, their conspicuousness, purely lexical, being both their means and their end. The purpose of pun is not to convey meaning, it is to convey double meaning (as in the quote by John Donne, where both meanings intertwine to stunning effect). I speculate that in the cases where the intention is to amuse, rather than to convey meaning, the attention is so much at the level of the lexical that the subtext of the emerging grammar string, while it is being 'fit' into the lexical arrangement, becomes slightly artificial. That is why subtext is not mentioned in the definition of wordplay, offered above.

The case of John Donne supports this assumption, in that his wordplay, perhaps not meant to amuse, but genuinely to arrive at two meanings that both fit the context, does contain subtextual meaning. 'Thou hast done' may be viewed as a grammatical string. It is contained in four contexts in the BNC (one of them negative and none religious), and in six contexts in COCA (half of them positive and could be described as religious, the others not religious and negative). If we go to the Google Books – UK corpus, however, we first encounter biblical contexts, which is appreciated because of the similarity of the context of situation. It turns out that 'thou hast done' can have a positive or negative meaning, depending on who is uttering the words. In the first twenty contexts, if God addresses man (or woman), they are (in these contexts) all negative, blaming mankind for various sins (unlike the context in COCA where God speaks to Abraham: '…because thou hast done this thing, and hast not withheld thy son, thy only son, that in blessing I will bless thee…). If it is man who speaks to God, then, naturally, or at least in the contexts studied, we witness only praise and trust in God's judgement. Thus, subtext in John Donne is in accordance with its contexts of situation, and, even if the roles of the speaker and hearer are reversed, it is human sins that come into focus, just like in Donne.

Larkin's 'and leaves the world alone' may serve to verify this assumption. It is perhaps less amenable to analysis, but at least the semantic aura is confirmed by the BNC.

The line ‘and *s the *’ yielded the following concordance (‘stray’ lines, containing grammar words, as well as those not containing verbs and nouns where they are found in Larkin’s line, and variants inappropriate for other reasons were excluded):

1 and has the power 12
2 and has the advantage 10
3 and takes the form 9
4 and opens the door 7
5 and forms the basis 6
6 and gives the impression 6
7 and has the potential 6
8 and reserves the right 5
9 and shuts the door 5
10 and concerns the land 4
11 and includes the PAT 4
12 and reduces the risk 4
13 and tells the story 4

Figure 7.6 The concordance of 'and *s the *'.

The most frequent QPVs, ‘power’ and ‘advantage’ do correspond to Larkin’s context, as the night is viewed both as powerful and an advantage from the point of view of the persona. Also, lines such as ‘has the potential’ and ‘reduces the risk’ in the concordance point to a positive semantic aura, although, of course, other contexts can be negative as well (for example, ‘opens the door’ is positive only in three contexts out of seven, and serves as a transition to a dangerous situation in the rest). Still, Larkin's line, like Donne's, contains subtext that fits the context of situation. The difference in the behaviour of subtext in these two poetic contexts and in the two-liner ‘Hot Potato’ is enormous.

7.5 Personification in a poem by Brodsky

As this section is dedicated to illustrating that semantic auras may be interconnected within a text, and that a short poem may be a good example of this because it is well thought-through by the author, this section will also include an example of personification deployed by the Russian poet Iosif Brodsky in one of his earlier works. It is useful to keep in mind Abrams's definition: Another figure related to metaphor is *personification*, or in the Greek term, *prosopopeia*, in which either an inanimate object or an abstract concept is spoken of as though it were endowed with life or with human attributes or feelings' (Abrams and Harpham 2009: 121).

The first half of the poem 'He knew that this pain in the shoulder' („*Он знал, что эта боль в плече...*") is considered in detail in Section 15, with reference to modal expressions Brodsky uses to pain his persona's attitude to cardiac pain that makes him die at the end of the poem. If at first he shrugs it off as a regular affair that will pass as usual, towards the middle of the poem he first suspects, and then realises, that he is in mortal danger – too late to summon help. Pain is the entity that is shown as taking on human characteristics, becoming the only interlocutor of a lonely man. To appreciate the context of situation, the reader ought to be aware that the Russian word *боль* (pain) is a noun of the feminine gender. This is the poem:

Он знал, что эта боль в плече
уймется к вечеру, и влез
на печку, где на кирпиче
остывшем примостился, без

движенья глядя из угла
в окошко, как закатный луч
касался снежного бугра
и хвойной лесопилки туч.

Но боль усиливалась. Грудь
кололо. Он вообразил,
что боль способна обмануть,
что, кажется, не хватит сил

ее перенести. Не столь
испуган, сколько удивлен,
он голову приподнял; боль
всегда учила жить, и он,

считавший: ежели сполна
что вытерпел -- снесет и впредь,
не мог представить, что она
его заставит умереть.

Но боли не хватило дня.
В доверчивости, чьи плоды
теперь он пожинал, виня
себя, он зачерпнул воды

и впился в телогрейку ртом.
Но так была остра игла,
что даже и на свете том
-- он чувствовал -- терзать могла.

Он августовский вспомнил день,
как сметывал высокий стог
в одной из ближних деревень,
и попытался, но не смог

названье выговорить вслух:
то был бы просто крик. А на
кого кричать, что свет потух,
что поднятая вверх копна

рассыплется сейчас, хотя
он умер. Только боль, себе
пристанища не находя,
металась по пустой избе.[37]

This is the concordance of *боль* (‘pain’) from Brodsky's poem:

[37] Source: http://www.world-art.ru/lyric/lyric.php?id=7521

1 хвойной лесопилки туч. Но боль **усиливалась. Грудь кололо.** Он вообразил,
2 кололо. Он **вообразил,** что боль **способна обмануть,** что, **кажется,** не хватит
3 влен,он голову приподнял; боль **всегда учила жить,** и он, считавший: ежели
4 его **заставит умереть.** Но боли **не хватило дня.** В доверчивости, чьи плоды
5 час, хотя он умер. Только боль,**себе пристанища не находя,металась** по пустой

Figure 7.7 The concordance of *боль* (‘pain’) in the poem by Brodsky „Он знал, что эта боль в плече...”

What follows is the overview of lexemes appearing in the main corpus of the RNC in the slot occupied by *боль* in the poem's patterns. The first line, * *усиливалась* (‘* intensified’) yielded the following results: *болезнь* (‘illness’) (7), *качка* (‘tossing of the waves’) (5), *буря* (‘storm at sea’) (4), *тревога* (‘anxiety’) (4), *боль* (‘pain’) (3), *гроза* (‘thunderstorm’) (3), *тоска* (‘anguish’) (3), *жара* (‘heat’) (2), *стрельба* (‘shooting’) (2), *темнота* (‘darkness’) (2).

These are all unpleasant sensations, weather conditions or circumstances. As far as the first line in the concordance of *боль* in Brosdky's poem goes, the word is used by the poet in accordance with the language norm. As for the second concordance line, the situation is beginning to change. The search line * *способна *уть* (‘* is able to *’) yielded the following results:

теория *способна сдвинуть*
сцена *способна вернуть*
соломинка *способна качнуть*
слава *способна вернуть*
свобода *способна вдохнуть*
правка *способна свернуть*
Россия *способна вернуть*
масса *способна примкнуть*
Катя *способна продрыхнуть*
женщина *способна упрекнуть*

The lexemes appearing in the wildcarded lexical slot are ‘theory’, ‘stage’, ‘straw’, ‘glory’, ‘freedom’, ‘editing’, ‘Russia’, ‘mob’, ‘Katia’, ‘woman’. Clearly, 'pain' is becoming a little more animate: there are six abstract inanimate nouns, followed by two groups of people, a feminine name and a ‘woman’. The inanimate ones predominate for the time being, with the ratio of six to four.

In the third search line, * *всегда учила *ить* (‘* always taught to *’) *всегда* (‘always’) was excluded because there were no matches. The search line * *учила *ить* (‘* taught to *’, the second wildcard in Russian having an infinitive ending) yielded three lexemes, all referring to women: *Наумовна*, *мать* и *маменька* (a patronymic, ‘mother’ and an old-fashioned form of ‘mommy’).

In the search line * *не хватило* (‘* ran out of’) only nouns in the dative case were taken into account, as in Brodsky’s text. Here are the first 20 lines in the main corpus of the RNC:

```
1 и вверх блестящие пузырьки. Скворцову не хватило дыхания, он вынырнул. Огляделся — Ли
2  в подобных случаях, казалось, что им не хватило времени для полного, совершенного ов
3                      К сожалению, мне не хватило времени, чтобы оформить документы, «
4                  Пусть некогда Лореасе не хватило сил, вдохновения и отрешенности, что
5    Разница в том, что генералу Бобырю не хватило элементарного человеческого мужества
6 е умные покупатели». Закрытие: Danone не хватило молока Danone останавливает работу к
7             А отчасти потому, что мне не хватило жизненного срока.
8 науки считают, что фашистскому режиму не хватило буквально месяцев для создания полно
9               [пахарь, nick] 2).клубу не хватило меда, и он пошел в тепленькую сторон
10 идел, глядя в окно. Естественно, ему не хватило времени. — Я должна тебе сказать одн
11                                 Кому не хватило места, везут в НИИ скорой помощи име
12                  [Анна С., nick] Мне не хватило этой сложной и противоречивой фигуры
13                     [AELEK, муж] Мне не хватило «его» в двух моментах.
14               Я получила пенсию, мне не хватило ее даже за квартиру заплатить.
15                     [АНС, nick] Всем не хватило внимания посмотреть, что написано в
16                           Потом Тасе не хватило карт для продолжения истории, и она
17                            Археологу не хватило бы и жизни на раскопку и инвентариз
18 ерской выносливости, которой Балтеру не хватило.
19               Расширенному Президиуму не хватило «аппаратного разума» Горбачева и его
20 от в чемпионате мира «до 18» Теймуру не хватило сил на финише, и он на очко отстал о
```

Figure 7.8 The concordance of * *не хватило* in the RNC

The lexemes in the wildcarded slots all refer to people, even if they are Danone (line 6), fascist regime (line 8), or the general party board (line 19). The exception is line 9, where the noun refers to a community of bees.

Brodsky's *себе пристанища не находя* (‘beside itself’, ‘moving restlessly about’) is an instance of word play, because the Russian original has the literal meaning of ‘could not find a place to settle’ (pain has nowhere to settle now the dying man is dead). In the RNC the agent is mostly animate, but abstract nouns referring to feelings, energy and similar, very occasionally also appear. The search line * *металась по* (‘* was thrashing

around') yielded 22 contexts, the wildcarded slot filled by two dogs, a cow, a female bear, a car, and 17 women of different ages in a state of extreme worry.

она металась по (22)
я металась по (5)
мать металась по (3)
Васса металась по (2; different sources*)*
бабочка металась по (2; different sources*)*

Personification in Brodsky is a matter of gradation. Its first mention is in line with the language norm. Then, predominantly inanimate, abstract nouns occupying the wildcarded lexical slot are gradually replaced by mentions of humans. Since the poem is written in Russian, and the nouns are feminine, the gender of these humans is female. Without entering into a debate as to whether pain is intentionally endowed by the poet with attributes of humanity or femininity, the corpus norm ought to have impact on the reader's understanding. The degree of such impact on particular readers would make the subject of a separate study. Even so, it follows from the text (given other clues that could not be included in this description) that not death as such, but solitary death is the greatest, though hidden, misfortune of Brodsky's persona (Милойкович in press).

On the basis of this extended example, it is possible to offer a tentative corpus-based definition of personification. It is a lexico-grammatical collocation in which the lexical slot is filled in the reference corpus with a noun denoting a human being. Shorter still, it is a lexico-grammatical collocation with human subtext. If we are also bold enough to attempt a corpus-based definition of gradation based on this one example, it can very tentatively be defined as a growing discrepancy between the text and the corpus norm in the direction that suits authorial intention.

If we compare personification, pun and metaphor at this stage of research, we might tentatively conclude that in pun the prevalent mechanism is *relexicalisation*, in personification it is *subtext*, and in metaphor it is *co-selection* of the implied, intuitively accessible meaning with existing context clues, in the presence of relexicalisation and rich,

multi-layered subtext. This makes metaphor not only the most complex and subtle, but also the most intentionally ambiguous and open to interpretation of the three devices.

7.6 Conclusion

This section has presented evidence in favour of the statement that semantic auras in a finished work of art are interrelated at the subconscious level, which can now be revealed using collocational analysis against the background of reference corpora. The deviation of lexico-grammatical combinations from their contextual norm not only reveals additional meanings, but these meanings often overlap and interact within a text. It could not be otherwise, because authorial intention must reasonably be manifested in any section of the text, from start to finish. The news is that now this subliminal content may be empirically extracted. This points to the great potential CPT must have for disciplines such as psychology and forensic expertise.

This section has also shown that collocation may reveal the mechanisms of figurative language. As previously stated, corpus-derived subtext is a pervasive layer of meaning in texts, and its unravelling always contributes to the text's interpretation. Still, a pun is a figure of speech which is sometimes intended to amuse. Therefore, grammatical auras embedded in it may not be relevant to the double meaning it conveys, if the double meaning is an end in itself, as is the case in comic one-liners. If, however, pun appears in serious texts, such as poetry, where both meanings may have their full realisations within the context of situation, subtext will be present and play its usual part, as a pervasive layer of meaning, within this text. For this reason, in the corpus-attested definition of pun, presented in this section, there is no mention of subtext.

In contrast, within Yeats's original metaphor, as it appears in the poem 'The Circus Animals' Desertion', subtext plays a very significant role. Both metaphors juxtaposed in Section 7.3 refer to the same intuitively accessible implied meaning (which can be described as 'literary devices', in the broad sense). Although the grammar strings in both metaphors are very different ('my * were all' vs. 'now that my * is' co-selected with 'gone'), the subtext in both is very similar. Not only is there a general aura of significance, but it contains, in both cases, two consistent groups of QPVs. One group which they share

denotes a significant person, like a family member, for example. The groups that they do not share are, however, very similar, because they can be described as 'cognition and spiritual life' in the first stanza, and an essential quality without which honourable existence is impossible, in the coda. The consistency of subtext in the different grammar strings containing a lexical item with the same or similar intuitively accessible implied meaning, as shown in Yeats's poem, confirms the ever-present role of collocation in meaning realisation, both in meaning creation and meaning construal.

According to Abrams and Harpham (2009: 121), personification is a figure of speech related to metaphor. In Brodsky's case, its subtext that is gradually constructed is the same as hidden subtextual meaning in Yeats: an unpleasant sensation turns into a vision of a woman. Our knowledge of the external world might suggest that in conditions more acceptable than those described by Brodsky it should be a woman, and not pain, restless and beside herself at the bedside of the deceased. This is the absence, subtextually recovered, that both Yeats and Brodsky share.

8 Forensic linguistics: diagnosing manipulative discourse

8.1 Introduction

This section reports on the first instance of applying Contextual Prosodic Theory to the area of forensic expertise with a view to illustrating how forensic linguistics may benefit by joining forces with CPT. In particular, such an application is a logical sequel to the research on insincerity, initiated in Louw (1993) and carried further in Louw and Milojkovic (2016), described here so far in Section 5.

Section 5.4.6 hypothesises on authorial insincerity detected at the level of corpus-derived subtext in Pushkin's line *передо мной явилась ты*. It is claimed in Section 5 that the poet may have inadvertently embedded in this line his reservations concerning the woman to whom his poem is dedicated. This type of insincerity is similar to the way the director of the British Council is quoted as using 'symptomatic of' while praising the University of Zimbabwe (Louw 1993: 170). There is no conscious intention to deceive on the part of the speaker – rather, he or she wishes to conceal some of their genuine attitude, perhaps from him- or herself as well as from the addressee. This, non-manipulative, type of insincerity must have first caught Louw's attention because Louw at the time was specifically interested in miscollocations as proof that semantic prosody was not accessible to intuition (see Section 2.2.2). As for manipulative discourse, Louw deals with it at the institutional level (Louw 2003, see Section 2.3), and mainly approaches it from the point of view of repeatable events: if the combination 'truth and reconciliation' was not recorded in reference corpora prior to the establishment of the Truth and Reconciliation Commission, then this established institution has no legal precedent.

This section (based on Милойкович in press 2019a) describes the first attempt at applying CPT as a rounded theory to an area in which purposefully manipulative language is one of the main focal points: linguistic expertise. The research question of this section is whether CPT may prove that the speaker/writer is intentionally lying, and whether CPT may contribute to existing non-corpus-informed methodology of conducting linguistic expertise – or, whether CPT-informed forensic linguistic analysis has any advantages over the non-corpus-informed kind.

8.2 Aims and methodology of the study

In order to illustrate the potential of CPT in the domain of linguistic expertise, I used a linguistic report of an expert investigation into a case of applying discrediting tactics.[38] The focus of the forensic investigation was an article published in one of Nizhny Novgorod newspapers, aimed at L., the chief executive of a major arms manufacturer and at the time one of the members of the City Council of a town in the Nizhny Novgorod region. The discussion section of the expert forensic report states that 'communicative discrediting tactics belongs to the group of uncooperative conflict-inducing communicative strategies, aiming at inflicting financial, moral or psychological damage on the addressee or the object of defamation, and is thus considered to be a speech act of verbal aggression' (*Иссерс* 2008). The aim of this section is to confirm, using corpus-stylistic methods, the existence in the article of the following attributes of verbal aggression singled out by Nizhny Novgorod forensic experts:

(a) termed 'content-related'

- numerous digressions and deviations from reporting the factual side of the event, containing irrelevant attitudinal, evaluative and expressive remarks on the personality and actions of the victim
- the creation of a false picture of reality in the direction that suits the needs of the aggressor, an intentional distortion of the 'world picture'
- an exaggerated and unfair account of the facts and events, omission of relevant detail, distortion of the factual side of the event in favour of emotional and evaluative statement
- irrational statements, alogisms

(b) termed 'language-related'

- amplification (overuse of negative expressions with a view to forming a certain impression in the addressee of the content that is being related)

[38]The text of the expert investigation was kindly provided by the Department of Forensic Expertise, Lobachevsky University of Nizhny Novgorod. I am particularly grateful to prof. T. B. Radbil' for his professional support in the matters of linguistic expertise.

- inadequate colouring of parts of the text or the whole text
- misplaced use of colloquialisms, jargon, or other language means that do not fit the general style of the register
- misplaced use of emotional and expressive constructions in the given context (Радбиль 2014; Радбиль, Юматов 2014).

Although experts divide these attributes of discrediting tactics into content-related and language-related, this section will show that they can all be diagnosed at the level of lexico-grammatical collocation within Firthian context of situation. The case of manipulative discourse that is being investigated is not a subtle one: the article contains expressions such as 'buffoon', which are easily classified as verbal aggression. Still, the existing forensic linguistic investigation conducted by a commission of experts will play a role in ascertaining how CPT may add to non-corpus-informed expert findings.

In accordance with the principles of CPT, the methodology used for this purpose will involve wildcarding and co-selection. The former will be used to establish semantic prosodies and corpus-derived subtext in the text of the article, and the latter will be of particular help in establishing states of affairs (Wittgenstein 1922) in the reference corpus and comparing them to authorial usages in the article. As for corpus-derived subtext, research (in particular Милойкович in press 2019a, on which this section is based) has shown that in Russian, due to the flexible word order of the language as compared to English, the method of considering lexico-grammatical collocations with the less frequent lexical variable wildcarded may be more effective than strictly following Louw's procedure for extracting subtext and wildcarding all lexical variables. The reference corpus consulted was the RNC.

8.3 Results of the study

8.3.1 The context of the study

The newspaper article, which is in the focus of the study and which was diagnosed by experts as a case of verbal aggression, was written around an incident in the town of A., caused by erroneous information posted by L. in social networks. L. wrongly claimed in his post that the famous actor S., who was then passing through the town of A., had been taken

seriously ill. According to the experts, the newspaper article exaggerates this single case into intentional circulating of false information in social networks, an occupation which L. supposedly pursues not only on a regular basis, but also during his working hours and at the state's expense.

This study will focus on the following excerpt from the article:

> *По словам супруги Щ-ва, похожий* ***случай*** *с ними однажды уже* ***произошел***. *Как-то Щ-в лег на плановое обследование в стационар.* ***Бульварная пресса поспешила распространить информацию, что*** *актер* ***якобы*** *при смерти. Но* ***одно дело, когда слухи распускают недобросовестные*** *журналисты, и совсем другое, когда это делает руководитель крупного оборонного предприятия – градообразующего завода А.! Это что, социально ответственный человек или* ***обыкновенный шалопай, сидящий в*** *соцсетях и* ***развлекающийся вбросами****? (****Очень наглядно такого рода личности обозначены*** *в сериале «Даешь молодежь» –* ***Кекс и Укроп****.)* ***Просто в голове не укладывается, зачем*** *директору солидного предприятия* ***сидеть в*** *соцсетях,* ***в готовности вбросить*** *туда информацию.* ***К тому же еще и недостоверную!*** (100 words, 46 highlighted (46%))

This is the English translation of the excerpt:

> According to the spouse of S., they had already had a similar experience when S. was undergoing a regular hospital check-up. The tabloid press hastened to spread the information that the actor was supposedly on his deathbed. But it is one thing when rumours are spread by unscrupulous journalists, and quite another when they are disseminated by the chief executive of a large state arms manufacturer supporting the whole town of A.! Is he a socially responsible citizen or an ordinary happy-go-lucky fool, sitting in social networks and entertaining himself with posting dubious information? (This sort of individual has been very clearly labelled in the serial *Hopeful Youths* as its personages Plumcake and Dill.) Why the director

of a sizeable factory should be sitting in social networks in readiness to post information, let alone of dubious nature, is simply beyond comprehension.

This paragraph from the newspaper article, in my view, well reflects the overall tone of the piece. It was chosen owing to its size and its place in the text. These are of relevance because, just prior to making these statements, the article quotes the statement to the press given by the actor's spouse, discussing the string of misunderstandings caused by the erroneous information which L. had posted online:

> ***Почему*** *о простудных болезнях артистов, о работе их внутренних органов* ***должен*** *знать весь мир?* ***Разве этично*** *такими вбросами о недомоганиях известного человека* ***поднимать рейтинг*** *своих интернет-ресурсов?* ***При этом*** *Л-в в сообщении назвал Щ-ва своим другом! Мне позвонила из Москвы помощница члена Совета Федерации Екатерины Лаховой, которая является другом нашей семьи, и спросила:* ***а что это у вас за*** *друг* ***такой появился*** *по фамилии Л-в?* ***Да какой он*** *друг? Мы его* ***даже знать-то не знали****. А теперь,* ***после того, что произошло****, могу с уверенностью сказать: Л-в —* ***неприличный человек****.* (87 words, 29 highlighted (33.33%))

This is the English translation of the paragraph:

> Why should a respiratory complaint of an actor be the property of the whole world? Is it ethical to post unverified information about a celebrity's health issues in order to raise one's rating online? On top of it all, in his post, L. calls S. a friend of his! I received a phone call from a friend of the family, Ekaterina Lakhova, who works as assistant to a member of the Council of the Federation. She asked, and what sort of friend have you acquired called L.? He is no friend of ours. We had never even met him. And now, after what happened, I can claim with assurance that Lavrichev is an indecent man.

These are two excerpts of comparable length, which appear one after the other in the article under investigation. The statement given by the actor's spouse will be hereafter

referred to as excerpt 1, and the excerpt which follows (with one sentence intervening) and which is written by the journalist himself will be referred to as excerpt 2. The segments in these excerpts highlighted in bold italics are those which were found, when studied using CPT methods, to have negative semantic prosodies and subtext.

The statement of the actor's spouse (excerpt 1), whose husband had fallen victim to rumours initiated by the post published by L., could have been expected to contain more negativity that the journalist's report (excerpt 2), as he had not been affected personally and as the incident he was describing could be described as relatively trivial in comparison with other issues normally discussed by newspapers. Nevertheless, the journalist's text, which should have retained neutrality and objectivity, was shown to contain 12.67% more negatively coloured words than the statement given by the agitated spouse of the famous actor. Indeed, the journalist could have shown a personal attitude to the incident (e.g. disapproval of the mechanisms of social networks), which might have explained the general negative tone of excerpt 2. Still, this is the conclusion given by the experts about this particular paragraph of the newspaper article:

> Here a one-time case of sharing unverified information is generalised without justification – allegedly, rumours, in the plural form, are disseminated routinely, as if this was a habitual action in the case of L. From this follows, surprisingly, a derogatory general statement describing the personality of L. as 'an ordinary happy-go-lucky fool, sitting in social networks and entertaining himself with posting dubious information'. Furthermore, we witness a negative belittling comparison with slow-witted personages of a TV show. This is a standard communicative mechanism of discreditation, when a single fact grows into an unjustified general negative conclusion about a person’s character. Such unfair exaggerations are normally signs that a person’s dignity is being intentionally depreciated.
>
> The efficacy of the strategy of discreditation depends on its serial nature, that is, when unfair exaggerations keep growing in scale. Gradation here is seen in the following indirect rhetorical question: 'Why the director of a sizeable factory

should be sitting in social networks in readiness to post information, let alone of dubious nature, is beyond comprehension.'

The effect of unfair exaggeration is enforced by yet more aggravated generalisation, which is aimed at creating a false image. The author is conditioning the reader to assume that L. continuously, all the time does nothing but spend time online posting dubious information (seemingly, at the expense of his professional duties). This is called amplification – intentional conglomerating of negative, distorted and exaggerated detail with a view to creating a negative background and tone, ultimately leading to misrepresenting the personality of the victim.

Both excerpt 1 and excerpt 2 will be analysed in the subsequent sections by means of CPT. This analysis will be juxtaposed with the conclusion of the committee of experts quoted above.

8.3.2 The analysis of excerpt 1

Let us consider the first sentence of excerpt 1. It is an indignant rhetorical question: ***Почему*** *о простудных болезнях артистов, о работе их внутренних органов* ***должен*** *знать весь мир?* ('Why should a respiratory complaint of an actor be the property of the whole world?') *Почему* ('why') was co-selected with the form *должен* ('should'), 1-5 words intervening[39]. The search yielded 1596 documents, 2282 entries. Here are the first ten documents:

1 ***Почему*** *в семье кто-то* ***должен*** *быть на привилегированном положении?*

2 Как по-вашему, ***почему*** *люди, которые* ***должны*** *по идее жить «на всю катушку», пробовать терпкий напиток бытия, встают на путь жертвы, становятся мучениками за идею?*

3 Не видно причин, ***почему*** *церковное священноначалие* ***должно*** *именно сейчас изменить свой подход и пойти на риск.*

4 Вряд ли нужно доказывать, что это несправедливо: природные ресурсы принадлежат всему обществу, и непонятно, ***почему*** *плоды их эксплуатации* ***должны*** *доставаться только его части.*

[39] The study was conducted in June 2017.

5 Глядя на эти преобразования, Киссинджер заявлял: «Я не понимаю, ***почему мы*** *должны стоять в стороне и наблюдать, как страна становится коммунистической из-за безответственности собственного народа».*

6a И ***почему****, кстати, именно пивовары* ***должны*** *платить за алкоголиков? Неужели вы, профессионал, тоже разделяете этот миф про «пивной алкоголизм»?*

6b Доказано, что это генетически обусловленное расстройство, угрожающее не всем поголовно, а приблизительно 10 процентам популяции. ***Почему*** *платить* ***должны*** *производители пива? Потому, что им дано право давать рекламу того продукта, который для этих 10 процентов популяции смертельно опасен.*

7 Это веская причина, ***почему*** *мы, на мой взгляд,* ***должны*** *бороться за должность губернатора.*

8 ***Почему*** *мильтон обязательно* ***должен*** *быть круглым болваном, глупее мальчишки — на этом строить не годится.*

9 Колюня сызмальства знал, что назвали его в честь деда, и честолюбивого старика это обстоятельство неимоверно тронуло, тем более его собственные сыновья, Колюнины дядья, назвали своих отпрысков иначе, но какое всё это имело отношение к Колюне и ***почему*** *он* ***должен*** *был улыбаться неприятному толстяку, не понимал и имени своего не любил, втайне мечтая его поменять и назваться, например, Виктором или Алексеем.*

10a Не хуже других… Почему я с одним Константином ***должен?!*** *Если он не может мне хорошую квартиру?!*

10b — ***Почему*** *мы вам свои деньги* ***должны*** *доверять? –* *запальчиво продолжала она.*

10c — ***Почему я должен*** *верить банку? — спросил Коноплянников.*

10d — Понимаете? — жить! ***Почему я должна*** *ждать? Мне нужно начинать ремонт!*

10 e И ***почему*** *я* ***должен*** *чувствовать себя виноватым?*

Figure 8.1 *Почему* ('why') co-selected with the form *должен* ('should'), 1-5 words intervening, main corpus of the RNC

Out of the 15 contexts, 6 (1, 10a, 10b, 10c, 10d, 10e) contain an indignant rhetorical question, of the type that is asked in excerpt 1. In 3 contexts (4, 5, 9) there is a lack of understanding and/or acceptance of the situation. These two categories of context are similar to the one in the opening sentence of excerpt 1. In yet another 3 contexts there is an expression of surprise at the illogicality of the state of affairs (2, 3, 6). In context 8 an author listens to criticism of a text he has written and submitted for publication. Generally, co-selecting *почему* ('why') and *должен* ('should') yielded contexts containing negative collocates (underlined in the concordance) - the only exception is context 7. In two thirds of

the contexts (in 10 out of 15) the speaker feels indignation and annoyance about the situation discussed. The rhetorical question in excerpt 1 particularly corresponds to the contexts yielded by the reference corpus in the sense that the reference corpus contexts for the most part describe trivial, everyday situations, except for contexts 4 and 7, which talk about issues of global significance.

In the expression *разве этично* ('is it ethical') the lexical variable was wildcarded. The search line *разве *но* yielded the following QPVs (in the English description QPVs other than adjectives have been omitted): *можно* ('permissible', 2013), *возможно* ('possible', 97), *трудно* ('difficult', 48), *нужно* ('necessary', 42), *случайно* ('accidental', 16), *важно* ('important', 16), *недостаточно* ('insufficient', 14), *обязательно* ('necessary', 13). There are altogether 149 **но* wordforms collocating with *разве*, but the collocate *этично* ('ethical') was not found. These QPVs constitute the corpus-derived subtext of the expression *разве этично,* such as 'is it permissible'.

The searchline *разве *но* yielded 3 355 contexts in 2 063 documents, in which the first 24 contexts were studied. They all showed that the answer to the question *разве *но* is always negative. For example, if the question is *разве можно?* ('is it permissible?'), the answer is that is it is not permissible. This is in accordance with the use of the expression in excerpt 1.

The collocation *поднимать рейтинг* ('raise the rating') gave 6 contexts in the main corpus of the RNC, 5 of which showed disapproval, e.g. *Но сейчас на носу выборы, и надо срочно* ***поднимать рейтинг*** *очередной чиновничьей партии* (‘But with the elections approaching, they must raise the ratings of yet another bureaucrat political party.’).

The expression *при этом* ('on top of it all') followed by an animate noun of the male gender (*при + это (loc1 / loc2) + S (nom, sg, m, anim)*) yielded 1788 contexts in 1404 documents. Out of the first 20 contexts, in 11 cases (55%) the persona referred to by the noun is given a critical account of, because of their inadequate speech or odd behaviour, e.g.: *При этом Шлиппенбах без конца занимал деньги у сослуживцев и возвращал их неохотно* ('On top of it all, Schlippenbah kept borrowing money off his co-workers and was never in a hurry to repay it.').

This is certainly the category of contexts that fits the usage in excerpt 1. There were also 4 contexts (20%) containing express approval, such as *При этом Спиваков ни разу не отменил ни одного концерта!* (On top of it all, Spivakov never cancelled a single concert!)

The expression *а что это у вас за* ('and what sort of * have you') yielded 10 contexts. In five the speaker is merely curious, and in five the speaker is suspicious. Clearly, the usage in excerpt 1 belongs to the latter category. The expression *что за * такой* ('what sort of *', a more frequent variation of the wording in excerpt 1) yielded the following QPVs: *вздор* ('nonsense', 23), *человек* ('man', 12), *народ* ('people', 8), *зверь* ('beast', 5), *черт* ('devil', 4), *шум* ('noise', 4). The QPV *вздор* ('nonsense') is in complete agreement with the usage in excerpt 1, as is shown by the first 32 contexts of *что за * такой* studied (overall 221 contexts in 184 documents were found).

Знать не знали ('had never met', 'had no idea') in the corpus is used negatively in 23 cases out of 27 (85.19%), in ways similar to the usage in excerpt 1: what the speaker had no idea about is negatively perceived. As to the expression *после того, что произошло* (‘after what happened’), it is used negatively in all the 44 contexts in the reference corpus, with varying degrees of negativity. This can of course be connected to the negative semantic prosody of 'happen', identified by Sinclair (1991: 112; see also Section 5.1.1). If one takes into account the consequences of the post as described in excerpt 1, the use of this expression is in accordance with the norm. In the corpus it usually refers to an event causing hardship or frustration.

The expression *неприличный человек* ('indecent man') was not found in the corpus - indeed, this sounds like a miscollocation, which the following findings will show. The adjective *неприличный* does not typically co-occur with a noun designating a human being. Out of the 85 collocations, only two contained a noun: *неприличный дворянин* ('indecent nobleman'), and *неприличный петрушка* ('indecent clown'). The most frequent collocates of the word form *неприличный* (an adjective of the masculine gender in the nominative case) are *жест* ('gesture', 15), *звук* ('sound', 7), *поступок* ('act', 6), *анекдот* ('joke', 5), *вид* ('look', 5), *вопрос* ('question', 4), *смысл* ('meaning', 4), *тон* ('tone', 4), *характер* ('character', 4). Indeed, *жест* (gesture) and *поступок* ('act') fit the context of excerpt 1

better than *человек* ('man'). Judging by the corpus data, this is the first miscollocation in excerpt 1, which can be explained by the agitated state of the speaker, as well as her wish to criticise L. Nevertheless, it is significant that the adjective collocates most frequently with *жест* ('gesture') and not at all with *человек* ('man').

All in all, excerpt 1 has been found to contain 33.33% of text carrying negative semantic auras. These usages all correspond to the norm as represented by the RNC, apart from the expression *после того, что произошло* ('after what happened'), which may also be taken as corresponding to the norm if it is understood to refer not only to the act of posting information online, but also to its consequences. The other usage that does not correspond to the norm is the non-idiomatic collocation *неприличный человек* ('indecent man'). These findings lead to the conclusion that the statement under investigation was issued in a state of indignation and annoyance, and that the speaker was not giving much thought to the consequences of her words.

8.3.3 The analysis of excerpt 2

Following Louw's (1993) notion of insincerity as an inadvertent breach of semantic prosody, the initial hypothesis concerning excerpt 2 was that the text, whose manipulative purpose could be easily detected without recourse to corpora, would turn out to contain miscollocations and breaches of semantic prosody as evidence of wishing to conceal the writer's genuine attitude. However, this did not occur. Collocations in excerpt 1 were frequent in the corpus and in this regard within the norm. At this point it became useful to consider the scale of prosodic and subtextual clashes published in Louw and Milojkovic (2016: 275) and its section termed 'uninspired insincerity'. It was claimed (ibid.: 308) that manipulative lies may turn into cliches. This section will present further evidence in favour of this assumption, which remained unproven in the book.

As previously stated, excerpt 2 contains 46% of negatively coloured text, which is 12.67% more than is contained in excerpt 1. The first negative SP in the passage is that of *бульварная пресса* (‘the tabloid press’). There were 9 occurrences in 6 documents in the main corpus of the RNC, and in 8 contexts out of 9 *бульварная пресса* co-occurred with negative collocates, as in: *Только шантажистские газеты да желтая* ***бульварная***

пресса *позволяют себе брать плату за разного рода статьи и заметки в тексте […] При чем же здесь прогрессивная, демократическая печать?* ('Only blackmailing newspapers and the yellow tabloid press go so far as to receive payment for printing various articles [...] How does progressive, democratic press come into this?')

Then, *поспешила *ить* (‘hurried to *’, the wildcarded slot is followed by an infinitive ending in Russian) is a characteristic example of a lexico-grammatical collocation, which should be preferred over its grammatical counterpart **ила *ить* ('*ed to *'). The latter corresponds to Louw's strategy for extracting subtext (Louw 2010b), but may prove to be too frequent in the corpus and a thankless task in practice. The RNC yielded 258 contexts in 191 documents. The first 30 contexts were studied. If contexts where *поспешила* ('hurried') refers to individuals (both in the sense of physical movement and of remarks or actions) are excluded, the rest may be classified into two groups: statements made or actions performed by politicians at the international or intranational level, and statements made by the press or a press service. The former may or may not be negatively evaluated, but the latter are always mentioned critically or ironically.

The sequence 'information that' (the noun in the accusative case) was contained in 72 contexts of the main corpus of the RNC. There were 7 positive contexts (9.72%), 47 negative ones (65.27%), and 18 were neutral or difficult to classify (25%). This is similar to the negative semantic prosody of 'give rise to', discussed by Xiao and McEnery (2006). However, the information is negatively viewed because of its untrustworthiness in only 9 contexts (12.5%).

Одно дело, когда ('it is one thing when') is contained in 143 contexts (131 documents) of the RNC, and has the semantic prosody of aggravation: one of the aspects of the context of situation was aggravated in 113 contexts (79.02%). A reverse tendency (improvement) was noticed in 27 contexts (18.88%). In the remaining three contexts (2.09%) the change does not involve either aggravation or improvement. It must also be stated that in the 30 contexts that went against the main semantic prosody, 17 (56.66%) described an initially unfavourable state of affairs.

If before the appearance of *одно дело, когда* ('it is one thing when') excerpt 2 was within the corpus norm when it comes to negative colouring, at this point the attributes of

verbal aggression set in, such as exaggeration, the creation of a false picture of reality, alogisms, amplification and inadequate colouring (see Section 8.2). Clearly, the journalist's use of the expression in excerpt 1 fits the main prosody in the corpus - aggravating an initially unfavourable situation. This is also the section of the text where the focus shifts from tabloid press to L. in particular:

*Но **одно дело, когда слухи распускают недобросовестные** журналисты, и совсем другое, когда это делает руководитель крупного оборонного предприятия – градообразующего завода А.!*

But it is one thing when rumours are spread by unscrupulous journalists, and quite another when they are disseminated by the chief executive of a large state arms manufacturer supporting the whole town of A.!

The aggravation aspect is not warranted by the context. As experts have stated, it was one erroneous post on the part of L. that caused the incident, and it occurred when L. was on holiday, and not in his role of 'a head executive of a large state arms manufacturer'. Moreover, as reference contexts showed, *одно дело, когда* ('it is one thing when') is normally followed by generalisation, the phrasing of a rule that works in all such cases. This aura of generalisation does not correspond to the single act of posting information committed by L.

More importantly, the reference corpus contexts show that the role of *одно дело, когда* ('it is one thing when') is to set the scene, establish the context of situation, describe the initial state of affairs that is later aggravated/improved/changed in other ways. For example, the initial state of affairs in the passage quoted below concerns retail prices of vodka:

***Одно дело, когда** настоящая водка стоит в провинции около 40 рублей, а «паленая» – около 20, при такой разнице большинство граждан не бросятся сознательно отравлять себя. А если, допустим, бутылка неподдельной водки при помощи акцизов дойдет до 5 долларов? Сколько у нас в год будет смертей от самоделок?*

It is one thing when a bottle of regular vodka costs around 40 rubles in the provinces, and of home-made 'burnt' vodka 20 rubles. With this difference in

prices the majority of citizens will not consciously poison themselves. But what if, shall we say, the price of a bottle of legally produced vodka is raised to five dollars owing to taxes? How many deaths a year will be caused by home-made beverages?

Here the initial context of situation revolves around prices on vodka. Then the fear is expressed that the situation may be aggravated by prices on vodka being raised, which may cause heavier consumption of dangerous home-made drinks. The reference corpus showed that this is a characteristic transition following *одно дело, когда* ('it is one thing when').

In excerpt 2 the initial state of affairs is rumours spread by unscrupulous journalists. Then the situation is aggravated through the change of participants in the context of situation - rumours are now spread by 'the chief executive of a large state arms manufacturer'. Thus L. is artificially incorporated into the context of situation as a participant who regularly, just like tabloid press, spreads rumours. As is noted by the forensic experts, this manoeuvre aims at creating the impression with the reading public that 'rumours, in the plural form, are dissiminated routinely, as if this was a habitual action.'

It may be taken for granted that *слухи распускают недобросовестные* ('rumours are spread by unscrupulous') and *обыкновенный шалопай* ('happy-go-lucky fool') are derogatory expressions. Still, it is of interest that in some reference corpus contexts *шалопай* is used condescendingly rather than disapprovingly, which also must contribute to the belittling effect noticed by the experts. Nevertheless, the most interesting remark of those in the concluding section of the forensic report I consider to be the following: 'The author is conditioning the reader to assume that L. continuously, all the time does nothing but spend time online posting dubious information (seemingly, at the expense of his professional duties)'. The statement which starts with 'seemingly' is a hunch, or a general impression formed on the basis of the imperfective aspect of certain verb forms in excerpt 2, such as *распускать* ('spread', 'disseminate'), *сидящий* ('sitting'), *развлекающийся* (entertaining himself), the expression *в готовности* ('in readiness to'), and generally as a result of repetitions in the passage. The experts used 'seemingly' because they cannot prove that such an implication (spreading rumours *instead of working*) exists in excerpt 2. They can only point to evidence for the implication that L. does so regularly. Still, the main

corpus of the RNC (8 contexts) and the newspaper corpus of the RNC (5 contexts) showed that co-occurrence of lemmas *сидеть* ('sit') and *развлекаться* ('entertain oneself'), with 1-5 words intervening, call up in the corpus those states of affairs where such activities stand in contrast to working, and not merely take place at a time other than working hours — there are 3 such contexts out of 5 in the newspaper corpus and 3 out of 8 in the main corpus. Moreover, all the contexts point to an activity that is performed while not working:

1 *Только и ходила в салоны красоты,* ***сидела*** *в ресторанах и* ***развлекалась*** *в клубах.*

2 А некий человек-паук, ***сидящий*** *дома за компьютером, искренне* ***развлекается****, зная, как сильно люди боятся ЕГО.*

3 *Готовиться к занятиям никто не мешает: хочешь — сиди и «ботай», хочешь —* ***развлекайся****.*

4 Влад Доронин прибыл без Наоми Кэмпбелл и ***сидел*** *букой, пока барышни* ***развлекались****.*

5 *Стояли последние дни щедрого лета, его ровесники загорали и купались в море,* ***сидели*** *в кафе,* ***развлекались*** *на дискотеках.*

Figure 8.2 Contexts of *сидеть* ('sit') and *развлекаться* ('entertain oneself'), 1-5 words intervening, newspaper corpus of the RNC

1 Как-то в очередной раз мы вчетвером ***сидели*** *при свечах за картами и без особого интереса* ***развлекались*** *этим делом*

2 Сидит *где-то дома, муж тем временем* ***развлекается*** *с любовницей, а потом явится и небось еще претензии будет предъявлять: почему обед холодный, рубашка не та выглажена…*

3 — Да вот, — продолжала хихикать П. П., — ***сидим*** *у моего приятеля, всемирно известного версификатора Канопуса, и* ***развлекаемся****…*

4 Сижу *и играю на гитаре, и не просто* ***развлекаюсь****, а посылаю призывные звуки в соседнюю аудиторию, где слушает какую-то лекцию моя будущая жена.*

5 ***Сидят*** *на газонах с семействами, вкушают,* ***развлекаются*** *с детьми, валяются в тени деревьев, — толпа красочная и шумная.*

6 Он молча ***сидел*** *или стоял, уставившись на что-либо, и по временам* ***развлекался****, дергая себя за свой хохол.*

7 Утомленные ожиданием люди ***сидели*** *и лежали на траве, тихо разговаривая, молодежь* ***развлекалась****, звенела гармошка, визгливыми голосами девки выкрикивали частушки.*

8 *Мы* ***сидели*** *в Чека – вдохновитель, быть может, «спекулировал» в Петербурге или* ***развлекался*** *в Москве.*

Figure 8.3 Contexts of *сидеть* ('sit') and *развлекаться* ('entertain oneself'), 1-5 words intervening, main corpus of the RNC

The next collocation, *очень наглядно* ('very clearly'), is mainly used negatively in the RNC (see Table 8.1):

Table 8.1. The use of *очень наглядно* ('very clearly') in the RNC

очень наглядно	**negative**	**positive**	**neutral**
main RNC (61)	25 (40.32%)	19 (30.64%)	17 (27.41%)
newspaper RNC (31)	20 (64.52%)	10 (32.26%)	1 (3.23%)

In any case, this expression emphasises how obvious a certain illustration is, which adds to the overall belittling effect of the text. The next phrase, *такого рода личности* ('this sort of individual') was found in the corpus only once, in a context authored by the famous Russian nineteenth-century satyrist M. Y. Saltykov-Shchedrin, and it was negatively used. The search line такой+род(gen/gen2,sg)+S(nom, pl) yielded 1445 contexts. Out of the first 25, there were 19 (76%) negative once, 5 (20%) were positive, and 1 (4%) was deemed neutral. In the following concordance the negative collocates are underlined:

1 ***Такого рода требования*** *не предъявляются к другим компаниям, поскольку для них предусмотрены иные механизмы контроля.*

2a ***Такого рода цифры*** *приводятся как аргумент, якобы подтверждающий полезность и эффективность общего надзора.*
2b ***Такого рода соображения*** *вызывают резкий протест в прокуратуре.*

3 ***Такого рода традиции*** *проявляют себя и в остальных областях духовной и материальной культуры.*

4 П роводите ли вы аналогии между конфликтами со «Спартаком» Цихмейструка и Сычева? – ***Такого рода дела*** *для нас – обыденное дело*

5 ***Такого рода предприятия*** *формировались и до послевоенных демократических реформ, но их учреждение приняло широкие масштабы позже, имея целью решение конкретных задач правительственной политики в финансовой, промышленной, научно-технической, социальной и других сферах.*

6a Я был уже далеко не мальчик, а более чем искушённый мужчина, был уверен, что где-то уже я и перегулял, и пресыщен, и уверен, что ***такого рода чувства*** *теперь не для меня.*
6b К министру культуры ведь обращаешься в чрезвычайных случаях — и при этом даже самые знаменитые артисты предпочитали, чтобы ***такого рода обращения*** *не происходили через голову руководителей Союзгосцирка, — а без «своего министра» мы и шагу не могли ступить.*

7 Это предварительное и общее определение сатиры, как и все ***такого рода определения****, неизбежно и абстрактно и бедно.*

8 Но он сам был когда-то разведчиком и прекрасно знал, что ***такого рода напутствия*** *к добру не приводят, — они расхолаживают даже самых верных своему долгу людей.*

9 Рука Кузьмина легла на телефон, и он собрался позвонить своему однокурснику Буре, чтобы спросить, что означают ***такого рода воробушки*** *в шестьдесят лет, да ещё когда вдруг кружится голова?*

10 Трое из гостей — Спешников, полковник и вице-губернатор, туповатый, приличный и скучный немец, — были ***такого рода люди****, что Вера положительно не знала, как их занимать и что с ними делать.*

11 Так вот, ***такого рода ситуации*** *дают нам возможность понять: раз ты остался жив, раз ты не оказался на вокзале или в этом троллейбусе, стало быть, тебе дали шанс.*

12 — ***Такого рода мегапроекты****, как «Сила Сибири» с ее ресурсной базой, как правило, не имеют прямой коммерческой окупаемости, разве что лет через тридцать и в основном за счет смежных эффектов при участии государства.*

13 ***Такого рода переживания*** *как чистилище: либо поднимают человека на уровень выше, либо ломают ему психику заодно с судьбой.*

14 Согласно докладу Минрегиона, ***такого рода инвестиции*** *в прошлом году выросли более чем в половине российских регионов.*

15 Государства, принимающие ***такого рода инвестиции****, пока не в состоянии конвертировать их в добавленную стоимость; при этом по обязательствам рано или поздно придется платить. Кризисы неплатежей будут конвертироваться в социальные кризисы.*

16 Если ***такого рода вопросы*** *окажутся в повестке выборов глав региона, резонанс будет едва ли не больше, чем от кавказских споров об этническом представительстве.*

17 Ранее ***такого рода замеры*** *делали только в отношении «горячего Юпитера», который обращается вокруг звезды Эпсилон Андромеды.*

18 Конечно, наверняка существуют неизвестные силы физического или квантового рода, препятствующие появлению ***такого рода червоточины.***

19 Пропуская через себя тонны горячей воды, ***такого рода станции*** *генерируют чистую электроэнергию, обеспечивая до 30 % потребления всего Камчатского региона и*

заметно снижая его зависимость от поставок мазута для «обычных» тепловых электростанций.

21 ***Такого рода злоупотребления*** *обрели самые разные формы – от торговли экспортно-импортными лицензиями и выгодными госзаказами до фиктивного банкротства частных предприятий, дабы продавать их за бесценок в частные руки.*

22 Что тут скажешь – неприятный день близится, ***такого рода новости*** *будут только прибывать, еще не один городок, надо полагать, вдруг опомнится и испугается (как заполошный Омутнинск), и еще не один юнец вздумает продавать места в заброшенном бункере (нашелся такой в Красноярске), и еще не один депутат отменит или запретит КС.*

23 ***Такого рода события*** *уже бывали, а потом оказывалось, что это статистические флуктуации.*

24 Взгляды варьируют от нигилистических, отрицающих сам факт возможности существования НЛО, до «упёрто-материалистических», склонных относить ***такого рода аномалии*** *к атмосферным явлениям или оптическим иллюзиям.*

Figure 8.4 The first 25 contexts of *такой+род(gen/gen2,sg)+S(nom, pl)* in the RNC.

Clearly, this collocation is used in excerpt 2 in an expressly negative meaning, which is how it is used in the most part of the concordance, pointing to a negative SP. Also, each context (with the exception of 6b and 10) refers to some sort of general rule established through practice or experience.

When it comes to the word form *обозначены* ('labelled'), in the RNC it is not normally used of people, unless it is a group, such as a cohort in social sciences. When it is used of people in other cases, it refers either to members of a persecuted minority (e.g. Jews), or to a criminal group (e.g. conmen). It is obvious that the context in which the word is used in excerpt 2 belongs to the latter group of contexts.

В голове не укладывается ('beyond comprehension') is another case of exaggeration, perhaps the most conspicuous in the passage. The 124 contexts in the RNC could be divided into three categories. The first came from the sports section and comprised commentaries on unsuccessful matches, manipulated scores and other unfortunate outcomes of sports competitions. After these 36 contexts were excluded, there were 88 contexts to be considered. Out of these, 23 contexts referred to murders, 9 to deaths and yet another 23 to shocking and deeply nontrivial events, including unethical behaviour - 62.5% overall. The murders were either accidental (e.g. killing a pedestrian as a consequence of

reckless driving) or brutal (e.g. accompanied by severe beating, rape, burning of the corpse). Other reported events which do not fall under these 62.5% are certainly less trivial than the error committed by L. described in excerpt 2. For example, a member of the Rostov Ballet company complains that his passport has been issued to another individual, therefore the dancer cannot renew his contract or retire: although this is not a brutal murder, the situation described is highly nontrivial.

Сидеть в ('be sitting'), co-occurring with *зачем* ('why', 1-5 words intervening) called up 39 contexts in the RNC. It is normally used critically of an action that is performed for no convincing reason, in the presence of more useful or enjoyable alternatives. As for the string *в готовности *ить* (in readiness to *), this too is an example of exaggeration. In the 26 contexts in the RNC, 12 (46.15%) are those of military action, in the literal sense. The rest describe firm intention, without negative connotations, bar one context:

Но они едины только в ненависти к Христу и ***в готовности губить*** *христиан.*

But the one thing they are unanimous in is their readiness to slaughter Christians.

In addition, *в готовности* * (in readiness to *) in the reference corpus shows the semantic aura of a permanent state. Neither firm intention nor permanent state, not to mention military action, have anything to do with the act of one-time posting of erroneous information online, which is the subject of excerpt 2. As for the Russian noun *вброс*, or verb *вбросить* ('dissemination', 'disseminate'), the reference corpus showed that such information is dubious at best.

К тому же еще и ('let alone') called up 144 contexts in the RNC, out of which 93 (64,58%) were negative, 33 (22,92%) were positive, and 15 (10,42%) were deemed neutral. This expression, used for emphasis, corresponds to the norm as found in the reference corpus. The information posted by L. online did indeed turn out to be false, although his intention to post false information has not been established.

All in all, excerpt 2 opens with negative semantic auras that are used within their contextual norm as established by the reference corpus. When the focus shifts to describing the incident of

posting false information, the negative semantic auras (apart from the expression *к тому же еще и* – 'let alone') cease to correspond to their corpus contextual norm and begin to call up (much) graver contexts than the situation described could provoke, e.g. *просто в голове не укладывается* ('beyond comprehension'), *в готовности* * ('in readiness to *'). Moreover, some turns of phrase disparage the status or personality of the victim (e.g. *обозначены* 'labelled'). Generally speaking, the mismanaged collocations tend to add not only the aspect of aggravation, but also of generalization, or that of a permanent state of affairs, e.g. *одно дело, когда* ('it is one thing when'), *такого рода личности* ('this sort of individual'), *в готовности* ('in readiness'). The aggravation and disparagement are enforced by the implication that the supposed perpetrator keeps causing harm during his working hours (sit+entartain oneself).

8.4 Conclusion

The section has described a study whose purpose was to illustrate how CPT could contribute to non-corpus-informed forensic linguistic investigation. This study continues research into insincerity started in Louw and Milojkovic (2016, Chapter 9), but this time the text chosen was diagnosed by Russian forensic experts as containing signs of verbal aggression and disparaging tactics.

Contrary to expectations, the section has shown that, in the piece of intentionally manipulative newspaper discourse studied, there were no observable breaches of semantic prosody or infrequent QPVs. Instead of miscollocations or infrequent combinations, the text contained ready-made phrases that are used in the reference corpus to describe far more aggravated situatrions. The clichés used also tended to add to the air of generalisation that pervaded the text under investigation. These characteristics were not detected in the statement to the press issued by a concerned party who had personally suffered from the actions of the criticised individual.

These findings show that a mere presence of negative auras cannot be interpreted as verbal aggression. Indeed, why can we not accuse the victim's spouse of adopting the strategy of discreditation? Her statement was only filled with negative auras, but also contained expressions such as 'indecent man'.

Even if she was not directly involved and suffering from emotional shock, we have empirical proof that her negative semantic auras were always within their contextual norm, apart from one expression (*после того, что произошло*, 'after what happened'). Her indignation was expressed within the scope of the triviality of her situation, which confirms the sincerity of her intentions. When it comes to the journalist, his text not only contained 12.67% more negativity, but his auras persistently and categorically breached their contextual norm in the direction of aggravation. This consistency, empirically proven at the level of grammatical strings, leads us to the conclusion that the misrepresentation, achieved by unnatural language use, was purposeful. A less cardinal mismatch would have led us to believe that some insincerity, conscious or not, could have taken place, but not with manipulative intention.

The data so far suggest that in cases of authorial insincerity mismatches occur within collocations (see Section 5), whereas, when it comes to manipulation, whole frequent collocations tend not to be used in accordance with their contextual norm. In any case, this section has shown that Louw's CPT may inform methods in forensic linguistics when it comes to proving manipulative intention.

9 Inspiration and Authorial (In)sincerity

9.1 Introduction

This section is based on Milojkovic's Chapters 9 and 10 of (Louw and Milojkovic 2016), which were begun in the attempt to define the mechanism of chunking 'events' at moments of inspiration. In the process of applying CPT to poets like Pushkin (Section 5), Yeats and Brodsky (Section 7), and Larkin (Milojkovic 2011a; 2012), it becomes obvious that a poet's wording often implies more than meets the eye. In other words, neither the reader nor the author can be consciously aware of all the implications of collocation in the authorial text, unless they are acquainted with corpus findings. Nevertheless, the poet/author has chosen that particular wording and not any other, given that there are many ways to convey approximately similar meaning.

Next, in all these cases the poem in question seems to have been inspired by a very private experience. The grammatical strings that were chosen for inspection seemed to convey the key meaning in both poems – in Pushkin's case, the fact of the woman's appearance, and in Larkin's, a comment on the power of love to erase past – wrong – encounters (see Section 4.5). It is a reasonably sound conclusion that the relevant strings were produced at moments of emotional fluctuation, which caused unexpected forms of subtext to surface in them. For example, when a person in love produces a verse that shows deviations from the norm that cannot have been consciously intended because they are only detectable computationally (Section 5), that verse must have been influenced by inspiration. Therefore, my hypothesis is that authors, especially poets, may show inadvertent insincerity not because they are 'at pains to conceal' (Louw 1993: 157), but because they are themselves unaware of the insight that is struggling out during the process of poetic creation.[40]

Also, it is often the case in poetry that the author is inspired to chunk one to several events, around which he or she then purposefully constructs the literary work and its world. Authors and poets have noticed that some sections of text were more inspired than others, not necessarily the initial verse(s). On the other hand, as a reader, I had sometimes thought

[40] I am grateful to Nenad Milojkovic for pointing out, in conversation, the connection between subtextual clashes and struggling insight at moments of inspiration.

I detected fluctuations in the quality of an author's writing that I thought might be attributed to a lack of, or to different degrees of, inspiration.

Therefore my initial hypothesis was that the mechanism of inspiration involved prosodic and subtextual clashes. The clashes themselves, as means to convey a subtle meaning, were thought to be inadvertently used, as semantic prosodies and subtext do not operate at a conscious level. However, according to Louw (1993, 2000), they will be indicative of deliberate irony in those cases where the poet, although unaware of the means, is aware of the meaning (see Herbert Williams's poem analysed Section 5.3). In those cases, however, where the poet is unaware of either the means or the meaning, the resulting clash is, according to Louw, indicative of unintentional insincerity. The hypothesis is that the inner conflict in the poet/writer manifests itself through a clash, as the poet/writer is unprepared consciously to face the insight he or she subconsciously expresses (see Milojkovic 2012 on Larkin).

In the previous paragraph I mentioned not only inadvertent authorial insincerity, but also authorial irony, as inspiration-induced. The hypothesis of the section states that inspired writing will differ from uninspired in the density of its subtextual and prosodic clashes, and that the clashes themselves will be indicative of the presence of inspiration. It is not relevant whether the author at the time of writing was aware of the meaning his or her text expresses – Lodge in 'bent on self-improvement' (Louw 2000) knew what he was saying, whereas Pushkin in *Передо мной явилась ты* ('you appeared before me', Section 5) did not, as is indicated by contextual clues. The cline offered in this section is an expression of this hypothesis.

It is possible to test this hypothesis additionally by contacting several poets and asking them to point out clear-cut cases of inspired writing, and that is what happens in Section 10: a British poet agrees to provide material for research. His comments seem to indicate that there are differences in the general understanding of the term 'inspired', applicable to different sections of a poem. This need not disturb the hypothesis that, generally, inspired text will contain more clashes than uninspired. Notwithstanding the limitations caused by authorial introspection, the manner in which strong sentiments are

converted into text at moments of high emotion make the effort of such analysis worthwhile.

It is hoped that an analysis of this sort will also help in establishing, tentatively, the connection between inspiration and quality of writing. Clichéd and banal writing will, presumably, be related to the absence of authorial irony or authorial insincerity as in Louw (1993), and generally by the absence of deviation from the norm that Louw refers to as devices (2008). At the other end of the cline there will be found texts classified as 'uninspired insincerity' – meaning not necessarily uninspired but 'at pains to conceal'. Forced writing, which comes not 'from the heart' but out of duty or any outward compulsion, will also be marked as 'uninspired insincerity'.

Finally, the question remains whether the word 'texts' used here refers only to written language. In truth, this question was never raised as the cline was meant to classify discourse generally. Arguably, forms of untruth might be more frequent in speech than in actual print (though forced or uninspired writing will occur), and clichés may not find their way into print much either, but will remain a conversational filler. Also, a person speaking off the cuff (e.g. the Director General of the British Council in Louw 1993) will be prone to slips of the tongue, unlike a poet who will shape thoughts and meanings carefully before letting them see the light of day. While nothing in the language is random, and I believe slips of the tongue to be as telling as purposeful writing, written texts are better material for at least the initial stage of this sort of study, precisely because they were consciously and intentionally composed.

In the studied passages, the strings that are the focus of research will be underlined, the lexical choices within them will be highlighted in italics, and the co-selected vocabulary and/or phrases which it forms will be highlighted in bold.

9.2 The cline of prosodic and subtextual clashes within contexts of situation

On a cline of the degree of deviation from the norm, the norm being dominant semantic prosodies and frequent quasi-propositional variables, the following division suggests itself:

CLICHÉ/ BANALITY/ SINCERITY/ INSPIRED SINCERITY/ INSPIRED INSINCERITY/ UNINSPIRED INSINCERITY

NON-

POETIC QUALITY DISCOURSE (POETIC, NON-POETIC) NON-POETIC DISCOURSE

DISCOURSE

Figure 9.1 The cline of prosodic/subtextual clashes within contexts of situation

The point of the cline is to present a flexible taxonomy based on the presence or absence of prosodic and subtextual clashes. The distinctions are not set in stone but are rather lines in the sand, as genuine texts will differ from one another in more ways than can be imagined or specified in advance. However, the Firthian context of situation will always prevent the blurring of the two major differentiating boundaries: the point of the cline is to delineate, tentatively and yet on the basis of textual evidence checked against the background of corpora, the distinction between quality and non-quality discourse.

9.3 Banality, sincerity, inspired sincerity in David Lodge's *Thinks* and *Changing Places*

Let us look at two excerpts from the novel *Thinks* by David Lodge. The following paragraphs refer to the major heroine of the novel. Still recovering from the shock of her beloved husband's sudden death, she finds he was routinely unfaithful to her:

> It seems that Martin had a *reputation* for *sleep*ing with his research assistants. Most of his close colleagues must have known this, including some whom I **met socially**. Many of the people at the Memorial Service must have known. Sandra herself had been in the church that day, though she avoided meeting me afterwards, and so, she assured me, had been the girl who supplanted **her** in Martin's **affections** (Lodge 2000: 201).

Passage 1

The use of grammar and vocabulary suggests no clashes, as the following data clearly show. In the underlined grammar string in the first line of the paragraph Lodge uses the most frequent quasi-propositional variable – 'reputation'. Given that the next QVP, ‘passion’, is twice less frequent, this finding works convincingly in favour of the hypothesis that this usage is completely non-deviant. The other QPVs also remarkably correspond to the context of the passage: ‘taste’, ‘gift’, ‘knack’, ‘penchant’, ‘talent’, ‘flair’, and even ‘job’:

1	had a reputation for	60
2	had a passion for	30
3	had a taste for	20
4	had a gift for	11
5	had a job for	11
6	had a knack for	10
7	had a penchant for	10
8	had a talent for	10
9	had a flair for	9
10	had a word for	9

Figure 9.2 The concordance of ‘had a * for’ from the BNC

The QPVs fit into the context and add different angles to the situation: a person who routinely committed adultery in the workplace must have liked it (‘passion’, ‘penchant’) and possessed the necessary ability (‘gift’, ‘talent’ and ‘flair’). Still, if we widen the string and search for ‘had a * for *ing), we will encounter similar QPVs (‘reputation’, ‘knack’, ‘liking’, motive’, mania’) in the long list of 105 concordance lines, but never will we find ’sleep’ occupying the second lexical slot. The Google Books corpora: One Million Books, UK and US are in complete agreement: ‘had a * for sleeping’ yielded no matches. Thus, although the author’s first lexical choice is the most frequent QPV, his second choice is unique. The impression of ‘ordinariness’ is deceptive.

'Must have known' is used 191 times in the British National Corpus and twice in this short text, in line 2 and line 3. The following are 11 random concordance lines from the BNC. The states of affairs that this string provokes are out of the ordinary and may involve blaming those who must have known, as is seen in Figure 9.3, where the relevant collocates are highlighted in bold:

1 course, I realize today, the place I hid them in **everybody must have known where they were** cos I was small and having to reach, and the large

2 in the **abuse** because they're allowing it to go on, they **must have known**, surely they must have known, **why didn't they do something**? In

3 in the **abuse** because they're allowing it to go on, they must have known, surely they must have known, **why didn't they do something**? In

4 suppose this is the work of antiques experts? (SP:PS68C) I think they **must have known** what they were coming for and this **I'm afraid**, is one of the

5 leaving me your notes on the making of the Big Glass when you **must have known perfectly well** that they contained material I would be **certain to find offensive**. This

6 you say about me and especially about my family, and which you **must have known** would give **offense**. However. Dear Harsnet, he wrote, just a brief

7 Ma was **always at her most unreasonable on do-days**, and I **must have known it was a do-day** because not only did we have extra help in the house

8 pseudonym, so cheques left on the mantel were **useless**, and he **must have known** that they would be. She had gone right to the bank with them when

9 the landlords or tenants for shoddiness and dirt? It seems obvious she **must have known** it was the fermer **but he wished she had said so, after all this**

10 apart. Even among perhaps a hundred girls in their year, they **must have known** each other quite well. Their ways had diverged at university -- Angela Morgan had

11 down towards the foot of the embankment. **Whoever had brought her here must have known** the place; you couldn't have picked it out in a hurry. The

Figure 9.3 The concordance of 'must have known' from the BNC

The atmosphere of tension prevails in all of the lines, except in lines 1 and 10. In line 2 (the same as line 3) 'must have known' is repeated twice, as in Lodge's text, and is followed by 'why didn't they do something?', which corresponds to the mood of the heroine of Passage 1. Line 4 is that of theft in a chapel, as is seen from the wider context, lines 5 and 6 mention offense. Line 7, describing a family matter, looks innocent in comparison to line 8 suggesting fraud. 'Her' in line 11 is a murdered girl, as the wider context shows:

> The girl had been missing for what -- a week? -- and off-hand he agreed with the local man's judgement that she had been here for most of it. The sapling trees had grown tall in the twenty-five years since the Beeching axe had fallen on the single railway track, and, even leafless as they were, they effectively screened the view from the top. A thick hedge shielded her from the field which swept down towards the foot of the embankment. Whoever had brought her here **must have known** the place; you couldn't have picked it out in a hurry. The country is not as deserted as all that, as McLeish, brought up in a village in Leicestershire, well knew.

There are 9 contexts of 'met' co-selected with 'socially' in the BNC. In Figure 9.4 they are given in full, in order to facilitate the reader's understanding of the contexts of situation. Surprisingly, the overall semantic aura is less innocent than one would expect from the combination of these two collocates. Contexts 1, 3 and 9 describe romance, and context 8 depicts a state of affairs that might lead to one. Religious debate is mentioned in context 6. However, the remaining four contexts introduce a sinister and even criminal tone: context 2 mentions death, context 4 prison, context 5 rape, and context 7 a court case. Context 3 depicts romance against the background of a police case, combining the two main auras of meaning connected with this collocation:

1 You rarely run into people you know in London; there are too many millions. But he and Eleanor moved in the same set **socially**. He **met** her at a literary dinner a couple of weeks later. He half wondered if she'd arranged it; women don't get invited to these things as often as men. They were not sitting next to each other, luckily, but he found himself opposite her accusing face over drinks and remembered the things he'd last said to her. He thought it best to give her a quick kiss and talk as if nothing had happened.

2 Freddie Mills and his boyfriend Michael Holiday each clasped Michael to them. Freddie Mills had been Michael's boyhood hero. It had been worship of him that had aroused Michael's interest in boxing. Nowadays he **met** him **socially**, as an equal. A few days before Anthony's death they had been together at the Lancaster Road Baths watching local semi-professional boxers.

3 She took his mind off Mandy, and that was strange considering she was on his daughter's case. He had no illusions about the police. He had been dealing with them on and off all his life. But Kate was the first plain clothes police officer he had ever dealt with on a personal basis. Oh, he had greased a few of the Old Bill's palms over the years, such as the Chief Constable's, but they were both in the Masons. Kate was the first member of the Force he had **met socially** because he wanted to. Because he enjoyed her company. She looked good enough to eat tonight. Red suited her. Her dark hair shone in the light from the candles. She looked softer somehow. More appealing. After all his empty-headed bimbos, he found he liked having a woman around him who demanded a bit of respect.

4 Only another one thousand and ninety four days, twenty three hours and --' So, what brings you here?' asked Pete, like they'd just **met socially**. as if they would move in the same circles. He probably thought a financial adviser was someone who told you to hand over the wallet or you'd get your throat cut.' You mean, what am --'' Yeah, what you in for?'' Embez --'' Fingers in the till, eh? Tut tut, a posh bloke like you an' all.' If Pete kept interrupting him, he could spend the next three years never finishing a sentence

5 He claimed to be a expert on child abuse and once even offered his services to the Metropolitan Police -- although we did not use him.' He was a doctor but that was about the only genuine thing about him.' In reality, we believe he was simply the perfect rapist.' One woman who **met** Courtney **socially** was sex therapist Sara Dale, 49 -- the Miss Whiplash evicted from Chancellor Norman Lamont's London home.'

6 Not only could they see, they could clearly be seen: every speaker could not but be aware of their non-Catholic auditors, and trim his words to cause as little upset as possible. And they talked regularly to, and presumably influenced, conciliar fathers they **met socially**. The observers found the proceedings difficult to follow, and had to be supplied with interpreters, for both pope and curia had insisted on Latin as the language of the debates

7 I hoped that Lord Lane would not prove of similar kidney. I had not **met** him **socially** before, but we had once clashed in court when he was counsel representing the police at the Brabin Inquiry into the Timothy Evans case. The first question he shot at me after the midday adjournment was,' Mr Kennedy, have you been discussing the case with your counsel during the luncheon break?'

8 They had **met socially** before the show went into rehearsal, seemingly all over Britain. The first run was at Stratford-upon-Avon, a place that was used to a different kind of theatre entirely, and ended at what was regarded as the graveyard of intimate revue, Blackpool. He made her laugh at those dinner parties and other gatherings where they had been together. On stage, and in the rehearsal room, he was something very different.' He was pretty terrifying.

9 When she left him in 1983 he was devastated. But after that he met up again with Olwen, a childhood friend. The two of them had often **met socially** in the old days, with their respective partners, at evenings in the Green Dragon, the local pub in the village of Welton, ten miles from Hull, where Horsley lived in a magnificent stone house which, he always stressed, did not have a drive. Olwen had had her troubles as well. Her marriage had broken up,

Figure 9.4 'Met' co-selected with 'socially' in the BNC

No one could have intuitively predicted the sinister element in 4 out of the 9 contexts of 'met socially'. Romance and the mention of death correspond to the content of Passage 1, although the criminal aura in the third of the contexts (4, 5, and 7) is unexpected: adultery is not a criminal offence. Generally, the contexts show that where 'met' and 'socially' do not chunk the state of affairs of romance, they tend to be used in criminal contexts (with the exception of context 3 mentioning both romance and a police case). If we accept that 'met' and 'socially' create a mixed aura of romance, tragedy and violence, the corpus data deepen the implication of the passage while not violating the overall message.

By co-selection we discover that 'supplant' does not collocate with 'affections' in the BNC or COCA, but fits into states of affairs caused by 'her' and 'affections', which are changeable and uncertain. The relevant collocates in Figure 9.5 are highlighted in bold:

1 ion that Mrs Yeo **was not secure** in **her** husband's **affections**,' Davidson said, primly. Mc
2 young French officer who **replaced** me in **her affections**.' He's trying to pass it off lig
3 in Adam's Creek, and Jed **won a place in her affections on his very first day by buying**
4 een them the family had **replaced** him in **her affections**. So now she asked if the young g
5 his success would mean **competition for her affections**? Didn't she realise how painful
6 e had managed to **supersede** Zacharias in **her affections**, and the three had dwindled to a
7 her. Amen. Grant that all carnal **affections may die** in **her**, and that all things belong
8 **not yet confident of any strong hold on her affections**, despite the fact that she seeme

9 To be reminded of the **distant object of her affections**. Well, what's so remarkable in t
10 l, was rather more **indiscriminate with her affections**.'' But you did become engaged?''
11 , " she said now, bathing them both in **her** overflowing **affections**, Gemma propped up on
12 uld see how he had **worked his way into her** aunt's **affections**. Great-Aunt Alicia had al
13 mitted shamefully. Guido **hadn't stolen her affections** from Arnie. There hadn't been mu
14 r.' The last I heard she'd **transferred her affections** to some aspiring West End produc
15 I'm the reason Matthew **isn't returning her affections**. She wants to believe that, if I
16 demanded, as if sensing **a challenger** for **her affections**.' Just a couple of times. As I
17 cy her, Vada **wonders where she stands in her** father's **affections**. She has to turn to T
18 nfully **trying to disinherit her son from her affections**, should feed his little toy so
19 nfully **trying to disinherit her son from** her **affections**, should feed his little toy so
20 nd dined by a romantic Italian **rival for her affections**. Pretending to be her husband,
21 d always **have come second to his work in her** husband's **affections** and that, because of
22 as she **becomes surer of her hold on the affections** of **her** poet, so she becomes surer o
23 as she **becomes surer of her hold on the affections** of her poet, so she becomes surer o
24 r for not giving her one. She **transfers her affections**, hitherto directed to her mothe
25 Mortimer, who **now commanded** the queen's **affections** and **her** actions, was the principal
26 she shows how **shallow** were the maternal **affections** that left **her** children to their fea
27 l,... taught from infancy **to subjugate** her passions and **affections**, **her** thoughts, her
28 l,... taught from infancy **to subjugate her** passions and **affections**, her thoughts, her
29 weak to carry on, but Rosa **replaced** her in my **affections**. **Her** Ladyship was a mistake r
30 weak to carry on, but Rosa **replaced her** in my **affections**. **Her** Ladyship was a mistake r
31 us experiencing penis envy and **transfers** her **affections** to **her** father as he has the pe
32 us experiencing penis envy and **transfers her affections** to her father as he has the pe
33 ourting Bella Wilfer, who **transfers** his **affections** to **her** younger sister Lavinia, to w
34 recognize that **another has claim upon her affections** -- father. What is the child to
35 l school, which has a special place in **her affections**; she watched its roll of pupils

Figure 9.5 ‘Her’ and ‘affections’ co-selected in the BNC

Although ‘supplant’ does not appear in the contexts, ‘replace’ appears 3 time (excluding the ‘doubled’ lines), ‘transfer’ appears 4 times (excluding the same), and ‘supersede’ appears once. The states of affairs show the semantic aura of transferability and even competition. Almost every line contains highlighted collocates confirming this conclusion.

On the basis of the searches performed in Passage 1 – one search for the subtext of a grammatical string, one for the semantic aura of another string, and two searches for the semantic auras chunked by two different pairs of collocates – one may reasonably conclude that the passage stays within the corpus norm with remarkable accuracy. This is supported by the subtext of the first searched string 'had a * for' becoming the author's choice, by 'must have known' chunking negative states of affairs, by 'met' and 'socially' revealing a sinister prosody alongside with that of romance, and by 'her' and 'affections' showing 'affections' to be constantly changing their object.

Despite the surprise of these discoveries, practically only one deviation from the norm has been revealed (in the second lexical slot of 'had a * for *ing'). The real surprise lies in the fact that the heroine ought to have felt violent and contradicting emotions during being told the painful truth, and that these emotional fluctuations ought to have surfaced during corpus searches in the form of deviations from the most frequent QPVs and states of affairs. Instead of this, we are presented with a passage conforming to the norm in a way that seems to be almost surreal.

The banality of the episode does not warrant the grief which the next lines tell us the heroine is feeling. 'I felt dizzy and hardly able to breathe. The raw breeze-block walls of the poky little office seemed to swell and contract, the gross, meaty nude on the Lucian Freud poster, and the black, burnished figure on the Mapplethorpe, seemed to ripple and move obscenely' (Lodge 2000: 201).

Furthermore, the paragraph is peculiarly focused on the social aspect of the discovery. The context of situation broadens to include, not only other mistresses as relevant persons, but also the circle of acquaintance who 'must have known'. Betrayal of this depth should not first invoke thoughts of one's standing in their social circle, unless the writer is making this one of the tell-tale signs of a character trait. The writer is actually expanding the context of situation to include others precisely in the attempt to convey the heroine's sense of shame and dismay, the sense that he has not conveyed otherwise. One might argue that the description of posthumous offense is purposefully banal, in deliberate contrast with the following description of shock. But what would be the point of deliberate banality? The offence is banal in a social context; in a personal context it is the opposite.

The assumption that the heroine was meant at this point to betray an overt dependence on her image in society is not supported by the rest of the novel. Further in the novel the heroine does not reveal other people's adulteries to the offended parties, so the implication that in the passage she resents not having been warned by those who 'must have known' is also ruled out.

All in all, the described paragraph is such a coherent match to the norm found in the reference corpora, both in terms of semantic prosody and in terms of subtext, that at times it becomes close to surreal – I am referring to the fact that 'had a * for' has 'reputation' as its by far most frequent variable, that 'met' and 'socially' are used together in death-related and romance-related contexts, and that 'her' and 'affections' combine to present the latter as changeable. None of these could have been predicted intuitively. Passage 1 is an example of ***sincerity*** verging on ***banality***, and thus its quality, according to the cline described above, is questionable. The other option is deliberate foregrounding on the part of the writer, by means of resorting to the banal where one would expect an understandable surge of emotion. Let the reader decide.

Passage 2 below also refers to the same incident in the novel:

> I *wonder*ed how long this philandering had been going on. As soon as I **pose**d the **question** I thought I knew the answer: since the time of my big depression, six or seven years ago. For six months I *languish*ed at the *bottom* of a deep **hole**, like the **shaft** of a waterless **well**, while kindly, puzzled people, of whom Martin was one, **peer**ed down at me over the **rim** of the **parapet** and tried to cheer me up, or lowered **drugs** and **advice** in a **bucket** (Lodge 2000: 202).
>
> Passage 2

The two underlined lines seem suitable for subtextual study, as carriers of potentially significant emotions, and also for practical reasons. 'I *ed how long' yielded 14 lines in the BNC, the only QPVs being 'wonder' (11), 'ask' (2) and 'enquire' (1). The same QPVs, in this order, were yielded by COCA and all Google Books corpora. Obviously, Lodge uses the most frequent QPV in the grammatical string which opens the passage.

The search string '*ed at the * of' also yielded results that confirm Lodge's adherence to the language norm, although his lexical choices do not correspond exactly to the most frequent QPVs. These are the most frequent QPVs from the BNC and the Google Books corpora:

Table 9.1 Results of '*ed at the *' in the BNC, the Google Books UK and the One Million Books corpora

	BNC	Google Books UK	Google Books: One Million Books
1	**died** at the **age** of (62)	**died** at the **age** of (30 038)	**died** at the **age** of (79 303)
2	**suffered** at the **hands** of (38)	**placed** at the **head** of (23 957)	**placed** at the **head** of (640176)
3	**achieved** at the **expense** of (26)	**placed** at the **disposal** of (23 957)	**placed** at the **disposal** of (43 018)
4	**displayed** at the **foot** of (24)	**placed** at the **end** of (10 010)	**suffered** at the **hands** of (21 708)
5	**issued** at the **end** of (22)	**suffered** at the **hands** of (9 705)	**placed** at the **end** of (21 019)
6	**retired** at the **end** of 22	**killed** at the **battle** of (8 588)	**received** at the **hands** of (19 965)
7	**announced** at the **end** of (19)	**received** at the **hands** of (7 463)	**killed** at the **battle** of (15 924)
8	**displayed** at the **top** of (19)	**printed** at the **end** of (7 243)	**situated** at the **foot** of (14 452)
9	**published** at the **end** of (19)	**situated** at the **foot** of (7 002)	**existed** at the **time** of (12 969)
10	**based** at the **university** of (15)	**mentioned** at the **beginning** of (5816)	**knocked** at the **door** of (12 756)

In the BNC, the two most frequent QPVs of this grammatical string are the negative 'died' and 'suffered' – Lodge's choice of the less frequent 'languished' fits in semantically with these. The third QPV is 'achieved', whose positive implication is undermined by the immediately following 'at the expense of'. The less frequent QPVs - 'displayed', 'issued', 'retired' 'announced' and 'published' are not attitudinally marked, but entail the use of the second QPV – which denotes an extreme point in either time ('at the end of') or space that can be described in height-related terms ('displayed at the top of' co-occurs with 'the page' in 17 cases, 'this screen' once, and 'the list' once). This corresponds to Lodge's use of 'the bottom' of the shaft. It must be noted that the string at position 11, 'excited at the prospect of' (not included in Table 9.1) flouts Lodge's negative use of '*ed at the * of', but it is much less frequent than the top two strings with 'die' and 'suffer' in the first lexical slot.

The Google Books UK results are slightly different, but the main prosody remains unchanged: 'died at the age of' is still at position 1, but 'suffered at the hands of' has

moved to position 5. The negative aura, however, is additionally supported by the string at position 6: ‘killed at the battle of’. The first positive use of the string is found at position 42 (‘married at the age of’), and the second at position 60 (delighted at the prospect of).

The Google Books: One Million Books corpus yields similar QPVs at similar positions, with ‘suffered at the hands of’ at position 4, and ‘killed at the battle of’ at position 6. ‘Married at the age of’ occupies position 30, ‘graduated at the University of’ is at position 37, and ‘delighted at the prospects of’ is at position 58.

These are co-selection (4 words to the left and 4 to the right) results of the passage:

- 'pos*' and 'question': 463 instances in the BNC and 1193 in COCA
- 'languish*'and 'bottom': 6 in the BNC and 4 in COCA
- 'bottom' and 'hole': 29 in the BNC and 210 in COCA
- 'hole' and 'shaft': 1 in the BNC and 10 in COCA
- 'shaft' and 'well': 1 in the BNC and 9 in COCA
- 'peer*' and 'rim': 6 in the BNC and 9 in COCA
- 'rim' and 'parapet': none
- 'drugs' and 'advice': 9 in the BNC and 4 in COCA
- 'advice' and 'bucket': none in the BNC and 1 in COCA

Although the analysis of two grammatical strings from the passage suggests no significant deviation from the norm, the co-selection results reveal that Lodge chunks less usual states of affairs through lexical collocation than in Passage 1. If ‘pos*’ and ‘question’, as well as ‘bottom’ and ‘hole’ are frequent occurrences, the rest of the collocations are considerably rarer. Out of the nine studied collocations, five are rarer than ‘met’ and ‘socially’ discussed earlier (9 occurrences in the BNC and 17 in COCA). One of the studied collocations, ‘rim’ and ‘parapet’, does not exist in any corpus mentioned in the Section 1.3; the least frequent one, ‘advice’ and ‘bucket’, is chunked through collocates that appear in different sentences:

a pot, some pan bread and a **bucket** of milk. **Advice** too: "Don't let this whip you, Rose. You

Frequency of occurrences as such is not a sufficient parameter of how rare a state of affairs is because we need to establish if these collocations are normally used in the writer's context of situation – namely, depression. 'To pose a/the question' is too frequent for such a search, so we may start with 'languish*' and 'bottom'. In the BNC, it is only used in the context of sports. In COCA, it is used in the contexts of a) the poor success of a TV drama, b) poor social status, c) a chain of priority of payment, and d) the position of a US state in the domain of providing education. Therefore, the collocation is used delexically in all these contexts, but never in the context of depression.

As for 'bottom' and 'hole', the contexts in the BNC are all literal but one exception, where both words are used delexically but do not belong to the same sentences or events. 'Hole' and 'shaft' chunk one context in the BNC where the collocation is used literally, and 10 contexts in COCA, where it is also used literally throughout, except for 2 contexts where 'shaft' is used in the meaning of 'beam of light'. 'Shaft' and 'well' are used literally in all contexts (1 in the BNC and 9 in COCA). The same goes for 'peer*' and 'rim' (6 in the BNC and 9 in COCA).

'Drugs' and 'advice' are used 9 times in the BNC (once in the context of depression), and 4 times in COCA (depression is not mentioned). Finally, distress, and not actual depression, is described in the only context of 'advice' and 'bucket' mentioned together.

To conclude, on the basis of the searches conducted it appears that no subtextual deviations in Passage 2 have been found. However, the states of affairs chunked in the passage are rare in the corpus and are almost never used in the context of depression. The absolute majority of the collocates under study are used literally in the reference corpora, while Lodge uses them delexically (starting with 'languished' and ending with 'bucket') to create a metaphor of the depressed state.

Another unusual state of affairs is chunked in Passage 3, describing one of the heroine's self-deprecating moods, during which she doubted her own quality as a writer:

> How often did Martin come home from work to find me long-faced and red-eyed because I'd lost faith in what I was writing. Once – it must have been before I had a

computer, or even a photocopier – he had to go out into the back garden and recover a whole manuscript from the dustbin where I had *thrown* it in a fit of despair, and brought it back **stained and smeared, but smelling rather pleasantly of the potato and apple peelings adhering to it** (Lodge 2000: 124).

Passage 3

The string 'where I had * it' called up the following QPVs in the four corpora. Although 'thrown' figures only at position 12 in the biggest corpus, Lodge's usage does not defy the most frequent QPVs. All data from all the corpora consulted have been included in Table 9.2:

Table 9.2 The QPVs of 'where I had * it'

	BNC	COCA	Google Books UK	Google Books One Million Books
1	found (1)	left (15)	left (482)	left (1246)
2	cut (1)	found (3)	seen (243)	seen (514)
3		dropped (3)	found (156)	found (407)
4		placed (2)	put (134)	placed (349)
5		tossed (2)	placed (133)	put (295)
6			got (101)	dropped (181)
7			heard (76)	laid (162)
8			dropped (67)	got (142)
9			laid (56)	heard (128)
10			hidden (50)	hidden (97)
11			hid (46)	entered (79)
12			lost (43)	thrown (73)
13				hid (67)
14				supposed (64)
15				concealed (63)
17				lost (63)
18				crossed (52)
19				taken (38)

Lodge's use is in accordance with those QPVs that share the meaning of 'put': 'concealed', 'dropped', 'hid', 'hidden', 'laid', 'left', 'placed', 'put', tossed', 'thrown'. In particular, it shares parts of meaning with 'left', which is the most frequent QPV in the three largest corpora of the four. This, and the QPVs 'tossed' (position 4 in COCA with 2 occurrences) and 'thrown' (position 12 in One Million Books with 73 occurrences) justify the conclusion that Lodge's lexical choice in 'where I had * it' is in accordance with the norm. There is no prosody of concealment in Lodge, which is present in the corpus findings. The act of placing the manuscript is emotionally coloured by the choice of 'thrown' without violating the language norm.

The string 'must have been before' called up 4 matches in the BNC, with 3 contexts clearly negative generally. The tendency is more specific in COCA, which yielded 5 matches:

1. before her death had been replicated
2. before I was born
3. before my time
4. before;
5. before my time

The overall implication of the contexts found in COCA is that the past 'before' a certain point is referred to as a happier time than the present (in the first four contexts out of five).

Studying the two chosen grammatical strings reveals that Lodge's usage does not violate the language norm, and deepens the reader's understanding. 'Thrown' is different from the most frequent QPV 'left' in that it is more emotional in the given context. 'Must have been before' adds the element of nostalgia for the happier times as opposed to the sadness/discontent of the present (COCA shows this aspect more definitely than the BNC).

As regards co-selection, mere intuition is sufficient to suggest that 'a manuscript smelling pleasantly of peelings' is an original state of affairs. However, what follows are findings based on reference corpora.

In the BNC there are no matching records of ‘manuscript’ co-selected with ‘smell’. COCA yields one co-occurrence of 'manuscript' and 'smell', where 'smell' is used delexically:

> Apparently, editors, interested only in making a dollar nowadays, just went to the web site of one of the on-line bookstores and checked your past sales rankings before they even read a **manuscript** - they could **smell** remainder fodder five city blocks away, and the writer's agent flatly told the writer as much.

These are co-selection results of the passage:

- 'back garden' and ‘recover*': no matches
- 'fit* of despair': 2 matches in the BNC and 9 matches in COCA
- 'stained and smeared': 1 match in COCA: ‘Big, glossy cookbooks are easy to give but pretty pointless to receive. Better to give the sort of cookbook that's going to get **stained and smeared** and dogeared because it's used so often.’
- 'stained' co-selected with 'smeared': no matches in the BNC, 1 in COCA: ‘Benjamin, hands black with oil, coveralls **stained**, cheek **smeared**, smiled again. His smile lit her heart like a match…’
- 'smell*' and 'pleasantly': 5 matches in the BNC and 17 in COCA
- 'pleasantly' and 'potato': no matches
- 'potato' and 'apple': 2 matches in the BNC, 21 in COCA
- 'apple' and 'peelings': none in the BNC, 2 in COCA
- 'peelings' and 'adher*' no matches

Co-selection results point to original collocations, some of which are not found in either the BNC or COCA. A person’s back garden is not associated with the verb ‘recover’, ‘pleasantly’ is not found in the proximity of ‘potato’, nor do peelings adhere in the reference corpora. Fits of despair are relatively rare (only 2 in the BNC and 9 in COCA). Also, all the collocates studied appear within the context of a manuscript, which adds to the oddity of the situation. ‘Manuscript’ collocates only with ‘recover* (once in the BNC and once in COCA), and with ‘adher*’ (once in COCA).

On the other hand, ‘stained’ and ‘smeared’ have produced two positive states of affairs in COCA, those of family cooking and human warmth.

While Passage 1 shows the least degree of deviation from the norm, Passage 2 (describing long-term depression) and Passage 3 (describing a fit of a writer's self-deprecation) apparently show no deviation from subtextual norms but chunk some unusual states of affairs. Both passages include lexical choices that are not normally employed within their respective contexts of situation (depression and a fit of self-deprecation). Passage 2 deploys this lexis to create a metaphor, and Passage 3 to re-create a past memory.

Passages 2 and 3 show Lodge as a writer whose style relies on creating unusual states of affairs in a way that engages interest – bearing in mind, of course, that the passages are written in a journal by a narrator, a female recently widowed writer, and the style may be characteristic of this particular character rather than of Lodge as a writer. Still, the difference between Passage 1 (by the same personage) and Passages 2 and 3 is conspicuous. While Passage 1 was described as *sincerity* verging on *banality* (this is still to be proved in this section), I will describe Passages 2 and 3 as ***inspired sincerity***. The term ***inspired*** is warranted by the presence of unusual states of affairs. The term ***sincerity*** is based on the lexical slots in Lodge's grammatical strings being filled with *frequent* quasi-propositional variables. In the future ***inspired*** will be used to denote unusual lexical collocation, or lexis that is not frequently used within a certain context of situation. The terms ***sincere*** and ***insincere*** will denote the presence or absence of subtextual clashes: instances where the choices of a grammatical string contradict the message of the text as shown by context clues. Thus, Lodge's 'bent on self-improvement' example will be described as 'inspired sincerity', while his 'life was what it appeared to be about' (see Section 4 of this section) will be termed 'inspired insincerity'. There is no distinction between the author and the narrator in this taxonomy, and no accusation that the author is 'insincere'.

In order to look at a different narrator in a novel by the same writer, let us consider a passage from *Changing Places*:

Eventually he had taken an apartment on the top floor of a huge old house owned by an Irish doctor and his extensive family. Dr O'Shea had converted the attic with his own hands for the use of an aged mother, and it was to the recent death of this relative, the doctor impressed upon him, that Morris owed the good fortune of finding such enviable accommodation vacant. Morris didn't see it as a selling point himself, but O'Shea seemed to think that the apartment's sentimental **associations were worth** at least an extra five dollars a week to an American **torn from the bosom of** his own family. He pointed out the armchair in which his mother had suffered her fatal seizure and, while bouncing on the mattress to demonstrate its resilience, **contrived at the same time to reflect** with a mournful sigh that it was scarcely a month since his beloved parent had **passed to her reward** from this very bed.

Morris took it because it was centrally heated – the first he had seen thus blessed (Lodge 1975: 57).

Passage 4

As in the previous examples, the underlined grammatical strings were chosen because, while carrying key moments in the passage, they are also 'searchable' in practice. Here are the most frequent QPVs from the four corpora for '*ed upon him':

Table 9.3 The most frequent QPVs of '*ed upon him'

	BNC	COCA	Google Books UK	Google Books: One Million Books
1	bestow (14)	bestow (33)	confer (41 823)	confer (108 597)
2	force (14)	force (21)	bestow (27 310)	impose (64 300)
3	impose (11)	descend (17)	impose (22 571)	bestow (64 261)
4	fix (8)	impose (16)	look (18 362)	look (47 390)
5	place (6)	impress (13)	call (17 620)	call (46 887)
6	look (6)	call (12)	force (16 262)	force (41 256)
7	confer (5)	look (12)	press (9 105)	fix (22 479)

8	descend (4)	place (10)	wait (8 159)	press (21 574)
9	call (4)	fix (10)	fix (8 1590	turn (20 423)
10	impress (4)	inflict (10)	pass (7 292)	wait (20 326
11	rest (4)			
12	turn (4)			

The most QPVs of this grammatical string are expressly positive: ‘bestow’ (in the BNC and COCA) and ‘confer’ (in the Google Books corpora). They are followed by expressly negative ones: ‘force’ and ‘impose’ in the BNC, ‘force’ in COCA, ‘impose’ in the Google Books corpora. In the One Million Books corpus ‘impose’ and ‘bestow’ are almost equally represented. Be that as it may, the most frequent QPVs betray a position of superiority on the part of the agent, not to mention the air of solemnity. Lodge’s choice, ‘impress’, is no. 10 on the BNC list and no. 5 on the COCA list. It is also no. 17 in Google Books UK and One Million Books. The semantic prosody of superiority and solemnity seems to contribute to the overprovision of the context of situation, as the comic figure of Dr O’Shea is very far from inspiring authority or awe; neither does the renting-an-attic-for-six-months context call for so much ceremony.

These are the QPVs of ‘thus *ed’ from the last sentence of the passage (in the case of the last three corpora the first ten QPVs are given):

Table 9.4 The OPVs of ‘thus *ed’

	BNC	COCA	Google Books UK	Google Books: One Million Books
1	thus **creat**ed (38)	thus **creat**ed (61)	thus **obtain**ed (88 412)	thus **obtain**ed (225 790)
2	thus **establish**ed (21)	thus **end**ed (43)	thus **describ**ed (64 247)	thus **form**ed (173 258)
3	thus **provid**ed (20)	thus **establish**ed (31)	thus **form**ed (60 099)	thus **describ**ed (140 128)
4	thus **form**ed (19)	thus **provid**ed (28)	thus **produc**ed (37 666)	thus **produc**ed (103 561)
5	thus **end**ed (18)	thus **help**ed (27)	thus **address**ed (27 686)	thus **prepar**ed (71 644)
6	thus **encourag**ed (16)	thus **produc**ed (26)	thus **prepar**ed (26 348)	thus **establish**ed (68 912)
7	thus **gain**ed (14)	thus **allow**ed (24)	thus **enabl**ed (25 103)	thus **enabl**ed (62 838)
8		thus **defin**ed (23)	thus **render**ed (24 564)	thus **engag**ed (62 210)
9		thus **requir**ed (23)	thus **express**ed (24 121)	thus **employ**ed (62 208)
10		thus **serv**ed (22)	thus **employ**ed (23 946)	thus **acquir**ed (61 651)

The quasi-propositional variables of this string do not include ‘blessed’ (it is not in the first hundred QPVs in the COCA and the Google corpora) but on the whole they share a constructive semantic aura. The exception is ‘ended’ (no. 5 in the BNC, no. 2 in COCA, no.2 in the Google Books UK and no. 49 in the Google Books One Million Books). The context of situation where ‘thus ended’ is used shows that it adds an ironic twist to the semantic aura of this grammatical string. This is the BNC concordance of ‘thus ended’ (the negative collocates are in bold):

1 aned forward to see what it was. # 15 # i. **Thus ended** the journey of Jekub, and the nom
2 is time in it as pleasant as possible, and **thus ended** up treating him far better than m
3 d, Warwickshire were playing the next day. **Thus ended** the most astonishing match the Co
4 bad that may **weaken or compromise** it. And **thus ended**, in effect, the doctrine of 'assimil
5 $30 million during the course of the year. **Thus ended** Jordan's **brief experiment in demo**
6 ted and the students and masters **expelled**. **Thus ended** Stamford's **prestigious university h**
7 ckers from 18 March 1975 to 25 March 1977. **Thus ended an era of electric street transpo**
8 condary School, and this was duly granted. **Thus ended** the Annual Examination, as the Bo
9 s of prohibition forbidding them to do so. **Thus ended** an episode of high drama, the exc
10 ", and **gave her hand to a brisker suitor**. **Thus ended** the Vicar's **sole excursion** into t
11 he legislature established by the USI who **thus ended** their own existence. As Vice-pres
12 ered out in 1934. The East Suffolk scheme **thus ended disappointingly** and in 1936 White
13 Orange landed with his forces at Torbay. **Thus ended** the **terrible tyranny** of the Stuar
14 l actor.' John Cassevetes got the part and **thus ended** Nicholson's **brief hope** of a majo
15 we **ended up with nothing for our trouble**. **Thus ended** Mary's wedding day. The party di
16 1536, when the occasion of the dispute was **thus ended**. Cranmer's greatest success in
17 ban-Delmas, **was beaten in the first round**: **thus ended** the period (1958-74) when Gaulli
18 bones, and got to bed at half past twelve. **Thus ended** the most remarkable day hitherto

Figure 9.6 The concordance of ‘thus ended’ from the BNC

Half of the concordance lines contain negative collocates of ‘thus ended’. Some lines (1, 2, 3, 8) are expressly positive. Generally, the concordance suggests that what ends in ‘thus ended’ is an episode of great significance. This contributes to the ironic effect of the passage, where the solemnity of the grammar is at odds with the banality of the context.

Two turns of phrase are underlined in the passage: ‘torn from the bosom of’ and ‘passed to her reward’. The person/personality who is supposed to be ‘torn from the bosom of his own family’ is Morris Zapp. The chunk ‘torn from the bosom of’ yielded 208 matches in Google Books UK and 549 matches in the Google Books: One Million Books

corpus. These are the 6 first random – and, as it happens, truly tragic – contexts found in the latter:

1 ... markets, and **bought by unfeeling tyrants**, they **sigh away in sullen despair a miserable life of hopeless degradation**. The husband is *torn from the bosom of* his wife, the brother from the embrace of his sister, and the shrieking babe from

2 ... the state in which, being found in the bosom, they are supposed to have been formed by the hands, of nature; anthropoplastic, the state into which, after having been *torn from the bosom of* nature, they have been brought by human labour.

3 Children were *torn from the bosom of* their families. The scions of our most noble I >\ ' houses were **transported into Siberia**, or were **compelled to enter the ranks of the debased soldiery** of Russia. The **Polish language was suppressed**, as well

4 'A king driven from his throne,' wrote Joanna Baillie, five years after the execution of Louis, 'will not move our sympathy so strongly, as a private man *torn from the bosom of* his family'.43 For the British, the Louis who **died under the guillotine**

5 Thus he said; and his companions obeying the implacable shade, the noble and unfortunate maid, and more than an ordinary woman, *torn from the bosom of* her mother, which she now cherished almost alone, **was led to the tomb**, and

6 She was *torn from the bosom of* her husband, and forced to become the mistress of the overseer. Thus was the **domestic felicity** of poor Jack **destroyed for ever** ; and he left with a thorn rankling in his bosom, to which there can be no parallel

Figure 9.7 Six first random contexts containing ‘torn from the bosom of’ in the Google Books: One Million Books corpus

The concordance speaks of the loss of personal happiness and physical freedom, and death. Morris Zapp’s voluntary departure from America for 6 months in order to continue teaching and research at a British university hardly warrants such pathos. The irony is deepened when the reader recalls the reason for Zapp’s departure: his wife Désirée threatened to divorce him because of, among other things, Zapp’s many adulteries.

The person/personality who 'passed to her reward' in the passage is Dr O'Shea's old mother. The phrase yielded 415 matches in the Google Books One Million Books corpus and 614 matches in Google Books US. These are the first six random contexts:

1 ... many others, of which she was an active and truly valuable member, Her loss is still deeply felt by us, and at every meeting, we painfully feel, that one to whom we looked as a guide and instructress, is no more, but has *passed to her reward*. (1834)

2 His dutiful wife and mother of his children, Bertha (Schriefer) Hoffmann, survived her worthy husband up to September 5, 1906, when she too *passed to her reward*. Issue : t. Henry, named Adelina Siemer, and of L Helena, their union they had ...

3 ... in Bayou Goula knew everyone else, Eugenie was the first among equals. Pete Berthelot recalled her death in 1980. "In early 1980, the greatest mind of Old Bayou Goula was silenced when Miss Eugenie (now 99) *passed to her reward* ... (2005)

4 In 1816, it was clear to her nuns that she was failing fast. With her sisters around her, chanting the Magnificat, she *passed to her reward* on April 8. She was beatified in 1906 and canonized in 1969. Blessed Thomas of Tolentino APRIL 9 From ... (1989)

5 With Mrs. Fielding *passed to her reward* with forgiveness of Henry on her lips,Jane and Henry are free to marry and consummate their love.15 Like the redeemed character he created, Charles Brockden Brown engaged in acts of radical ... (2012)

6 When she *passed to her reward* 9 April 1889 in her seventy-second year, all of her eleven children had accepted the faith. Father died in Peru, Kansas, February 1904. : Winter weather was too severe in northern Indiana. Usually I suffered ... (1997)

Figure 9.8 Six first random contexts containing 'passed to her reward' in the Google Books: One Million Books corpus

While ‘torn from the bosom of’ calls up contexts of cruelly broken lives, ‘passed to her reward’ is in the reference corpus used of devout matriarchs and valuable members of religious communities.

As for states of affairs chunked by co-occurrence, the chosen chunks are ‘associations were worth (at least an extra five dollars a week)’ and ‘contrived to reflect’. Both combinations, ‘associations’ and ‘worth’, and ‘contrived’ and ‘reflect’ sound unusual and humorous without recourse to corpora. However, these are corpus data:

- a 'association*', 'worth': 1 match in BNC and no matches in the COCA:

around of thirty thousand pounds. Even the **Association** headquarters was **worth** less than its mortgage. He agrees that the Conservative party

- 'contriv*', 'reflect*': no matches in the BNC and 1 match in COCA:

and reading them resembled reading a collection of stories, **not** a collection **contrived** to educate or **reflect** social need, **but an honest collection**, filled with voices

Never in any reference corpus are a person’s associations evaluated in cash. As for the only example where ‘contriv*’ and ‘reflect’ appear together, this example suggests that an attempt ‘to contrive to educate or reflect’ is the opposite of what its author considers honest.

In this passage, Lodge masterfully conveys Zapp’s impression of Dr O’Shea. He is a silly fool, pompous and petty, but nowhere in the passage will we read this qualification. This impression was conveyed to the reader by the following means:

- original states of affairs (e.g. 'associations were worth', 'contrived to reflect')
- fractured prosodies leading to overprovision, i.e. the prosodic clash of 'torn from...' (violent and permanent breaking of families) and the semantic prosody of 'passed to her reward' (a devout matriarch and a key member of a religious community)
- subtext (Lodge’s lexical choices in grammatical strings are infrequent QPVs but fit in with the pattern shown by the reference corpora)

These characteristics make Passage 4 an example of ***inspired sincerity***.

On the cline, Passage 4 is further away from *banality* and closer to *inspired insincerity* than Passages 2 and 3 described above, as the QPVs used in the grammatical strings are much less frequent – and also because the writer has employed fractured semantic prosodies, which was not the case in Passages 2 and 3. On the basis of these differences, I consider Passage 4 ***more inspired*** than Passages 2 and 3. I do not consider it more insincere as there are no subtextual clashes in the passage (i.e. the lexical choices in the grammatical strings fit into the QPVs found in the reference corpora).

Here it is suitable to raise the question of devices. CPT offers a modification of the view that devices are commonly and consciously recognized by the author and the reader alike. Let us take 'For months I languished at the bottom of a deep hole' from Passage 2. 'I languished at the bottom of a deep hole' is a metaphor. It contains no fractured prosodies – 'I' and 'languished', 'languished' and 'at the bottom', etc, are not unexpected combinations. The metaphor creates underprovided context of situation and clashes with the surrounding context (writers do not languish in holes in the midst of civilized Britain), and here lies the deviation from the norm. The deviation is created by means of the context of situation, and can be recognized without recourse to corpora.

When Lodge says that conference goers are 'bent on self-improvement' (Lodge 1984: 2), the collocational mismatch is unconscious and cannot be recovered without corpora. There are no collocational mismatches in 'languished at the bottom of a deep hole'. The metaphor and the irony are both devices, but they differ in respect of fractured semantic prosodies. The hypothesis of this section is that the author was *more inspired* when he came up with the irony, because the mismatch cannot have been conscious. The other hypothesis is that the metaphor is more predictable than the irony, and that *the very unpredictability of fractures of prosodies is a sign of the author's greater inspiredness*.

Having illustrated how the level of inspiredness of the text may vary depending on the presence and absence of fractured prosodies and unusual states of affairs, let us look at the distinction between sincerity and banality in the writing of the same author – since the claim was made that the very first analysed passage, Passage 1, may be accused of banality.

For this purpose, a sincere passage from the very ending of *Changing Places* will be analysed:

> As you are reading, you're aware of the fact that there's only a page or two left in the book, and you get ready to close it. But with a film there's no *way* of *tell*ing, especially nowadays, when films are much more *loose*ly *structure*d, much more ambivalent, than they used to be. There's no *way* of *tell*ing which frame is going to be the last. The film is going along, just as life goes along, **people are behaving**, doing things, drinking, talking, and we're watching them, and at any point the director chooses, without warning, without anything being *resolve*d, or explained, or wound up, it can just… *end.*
>
> PHILIP shrugs. The camera stops, freezing him in mid-gesture. (Lodge 1975: 251)

Passage 5

While practically every grammatical string can be searched for its subtext, the underlined strings were chosen both because they carry key transitions in the text, and because they are suitable for practical reasons.

In the spirit of experimenting with various forms, which pervades *Changing Places*, the ending is written in the form of a film script. 'There is no way of telling' appears in the passage twice: Swallow explains that, unlike when reading a book, a person cannot tell when the film they are watching is going to end. 'There is no * of *ing' is a 6-word string and can only be searched in the BNC and the COCA (the Google corpora only offer the possibility of searching strings consisting up to five words). However, the results yielded are sufficient to prove that Lodge's choice does not deviate from the string's subtext. The top five most frequent QPVs are provided in Table 9.5 below:

Table 9.5 The frequent QPVs of ‘there is no * of *ing’ in the BNC and COCA

	BNC	COCA
1	there is no **way** of **know**ing (27)	there is no **way** of **know**ing (53)
2	there is no **way** of **tell**ing (13)	there is no **hope** of **find**ing (4)
3	there is no **way** of **decid**ing (3)	there is no **possibility** of **gett**ing (3)
4	there is no **chance** of **gett**ing (3)	there is no **way** of **mak**ing (3)
5	there is no **way** of **ensur**ing (3)	there is no **way** of **tell**ing (3)

Bearing in mind that in Philip Swallow’s speech ‘telling’ may be replaced by ‘knowing’ without a change of meaning, it is clear that Lodge’s use of the string is conventional: ‘way’ and ‘knowing’ are the most frequent QPVs. ‘Chance’ and ‘possibility’ in the first lexical slot deepen the implication of the string (a person would like to know when a film is going to end). The QPVs in the second slot that are not ‘know’ or ‘tell’ are ‘decide’, ‘get’ and ‘ensure’ in the BNC, and ‘find’, get’ and ‘make’ in the COCA. They underline how desirable that knowledge would be.

‘Films are much more loosely structured,’ says Swallow in the passage. The string ‘are much more *ly *ed’ consists of five words. One gets the feeling that the BNC and the COCA are not large enough to present a valid sample of the string’s usage. The BNC contains 10 different ‘*ly *ed’ combinations following ‘are much more’: ‘strongly sclerotized’, ‘richly tracheated’, ‘reasonably priced’, ‘readily separated’, intimately mixed’, ‘frequently applied’, ‘directly related’, ‘deeply embedded’, ‘closely intertwined’, and ‘carefully controlled’. All these appear only once. It would seem that one occurrence from the BNC – ‘readily separated’ – corresponds to Swallow’s ‘loosely structured’, whereas three occurrences – ‘intimately mixed’, directly related’ and ‘closely intertwined’ – oppose it. The rest of the occurrences of this string are not relevant to Swallow’s speech.

As for the COCA, it contains 35 diverse and singular occurrences, and two collocations that appear twice: ‘are much more closely tied’ and ‘are much more closely related’. These appear to contradict Swallow’s ‘loosely structured’. Still, given that both the BNC and the COCA mostly provide single occurrences with varying QPVs, the much

larger Google Books corpora may be more reliable when it comes to the generalisation of findings.

Table 9.6 presents ten most frequent occurrences of the string ‘are much more *ly *ed’. ‘Closely related’ is at position 1 as by far the most frequent occurrence:

Table 9.6 The most frequent QPVs of ‘are much more *ly *ed’

	Google Books UK	Google Books One Million Books
1	are much more **close**ly **relat**ed (225)	are much more **clos**ely **relat**ed (792)
2	are much more **wide**ly **distribut**ed (153)	are much more **frequent**ly **affect**ed (450)
3	are much more **high**ly **develop**ed (119)	are much more **high**ly **develop**ed (376)
4	are much more **strong**ly **mark**ed (118)	are much more **wid**ely **distribut**ed (315)
5	are much more **frequent**ly **affect**ed (99)	are much more **strong**ly **mark**ed (257)
6	are much more **common**ly **us**ed (94)	are much more **nearl**y **alli**ed (216)
7	are much more **close**ly **connect**ed (86)	are much more **close**ly **connect**ed (208)
8	are much more **wide**ly **us**ed (81)	are much more **close**ly **alli**ed (207)
9	are much more **near**ly **alli**ed (65)	are much more **widely us**ed (194)
10	are much more **close**ly **link**ed (65)	are much more **common**ly **us**ed (170)

In meaning, ‘closely related’ seems to contradict Swallow’s ‘loosely structured’. This is in agreement with the BNC and the COCA findings, where relevant collocations (those that could be related to the context of Passage 5) were in disagreement with ‘loosely structured’ rather than supported it. The other ‘supportive’ QPVs are ‘closely connected’ (position 7 in both corpora) and ‘closely linked’ (position 10 in Google Books UK). The first lexical slot in the studied string is never occupied by ‘loosely’ in any corpus. The second lexical slot is never occupied by ‘structured’.

Interestingly, position 11 in One Million Books is occupied by ‘are much more widely separated’, which contradicts the most frequent QPVs in both Google Books corpora. The same collocation occupies position 17 in Google Books UK and is followed by ‘are much more widely spaced’ at position 18. In the Google Books texts featuring QPVs ‘closely related’ (no. 1) and ‘widely separated’ (no. 11), these are apparently mostly

used in strictly scientific contexts, in which 'separated' denotes a looser relationship. In the context of the most frequent QPVs of the string, Lodge's 'loosely structured' is not out of place – it fits semantically, as it answers the question of *how* two entities are *related* (as in 1, 7, 9, 10 of the Google Books UK concordance, and 1, 4, 6, 7, 8 of the One Million Books). A film is more loosely structured than a book, according to the passage, and that is an aspect of their relationship. Therefore, although Lodge does not use frequent QVPs in the second discussed string, his choices are not semantically different.

Books and films are about life; films and life end unexpectedly, according to the passage. 'The film is going along, just as life goes along[...] and at any point[...] without anything being resolved [...] it can just… end,' says Philip.

The string 'without anything being *ed' yielded no matches in the BNC and 3 matches in the COCA ('waste', 'move' and 'explain') as QPVs, each occurring once. 'Without anything being explained' fits the context of the passage while deepening its interpretation – at the end of a film we expect a denouement, which will include the explanation of whatever motives, actions and events that have hitherto remained unclear. The whole Google Books UK corpus offers 'settle' as the only QPV, appearing 58 times. This is a perfect match to Philip's words, as 'resolved' and 'settled' are very similar in meaning in the given context of situation. Google Books One Million Books yield the following QPVs:

1. without anything being accomplished (143)
2. without anything being settled (118)
3. without anything being added (113)
4. without anything being discovered (52)
5. without anything being decided (42)

The QPVs 'settled' and 'decided' correspond to Philip's 'resolved', while the most frequent 'accomplished' and the much less frequent 'discovered' add to the interpretation of Swallow's sentiment: a film, as well as a life may end without anything resolved, settled, accomplished or discovered by its main protagonist. Given the fact that in the largest corpora 'settle' figures the most prominently (it is the only QPV in Google Books UK and

the second most frequent in One Million Books), Lodge's use of 'resolve' may be considered to fit the norm.

Both a film and a life 'can just... end', says Philip. The grammatical string 'it can just *' yielded very little material in the reference corpora. The Google Books UK and One Million Books produced 'it can just as' (483 and 973 occurrences) and 'it can just be' (339 and 506 occurrences). The former is a grammatical string in its own right and may be ignored. More diverse variables were produced by the BNC and the COCA:

Table 9.7 The variables in 'it can just *' in the BNC and COCA

	BNC	COCA
1	it can just as (6)	it can just be (15)
2	it can just happen (2)	it can just as (9)
3	it can just strike (1)	it can just build (2)
4	it can just add (1)	it can just get (2)
5	it can just go (1)	it can just hang (2)
6	it can just distinguish (1)	it can just keep (2)
7	it can just be (1)	

It appears that, overall, the most frequent collocation in the reference corpora is 'it can just be'. Next comes 'it can just happen' in the BNC, with only two occurrences, and with the well-known negative semantic prosody of 'happen' (Sinclair 1991: 112). Lodge's choice of 'end' here appears as something that is simply a part of life, as 'be' and 'happen' indicate. The QPV 'strike' is also significant as it adds to the negative view of endings that cannot be predicted, except for the known fact that sooner or later they will inevitably happen/strike.

Subtextually, the passage is non-deviant. Lodge does not use the most frequent QPVs in every grammatical string; however, most QPVs are semantically related to the frequent ones. In the first analysed grammatical string, 'there is no * of *ing', Lodge uses the most frequent QPV, 'way', in the first slot, and in the second slot he uses a synonym ('tell') for the most frequent QPV in this slot ('know'). The second analysed string, 'are much more *ly *ed', is frequently used to describe how two entities may be compared. Although Lodge's choices, 'loosely' and 'structured', are not found in the corpus, they semantically fit into the string's general use. However, every reference corpus yielded more

combinations which contradict ‘loosely structured’ than those which support it. The most frequent QVPs in the COCA and the Google Books are ‘closely related’. This means that Lodge’s use, although within the norm, is not the most frequent.

In the third analysed grammatical string, ‘without anything being *ed’, Lodge uses the lexical item ‘resolve’, which fits perfectly with the most prominent QVP – ‘settle’. As for the fifth grammatical string, the QPVs ‘be’ (no. 1 in the COCA and no. 2 in both Google Books corpora) and ‘happen’ and ‘strike’ (no. 2 and no. 3 in the BNC) are significant. As the most frequent lexical QPVs, they show the universal and potentially unwelcome aspect of endings.

As for lexical collocation, the only noticeable deviation from the norm is perhaps the chunk ‘people are behaving, doing things, eating, talking’. ‘People are behaving’ is obviously never used without additional information as to what people and/or behaving how and/or where, in what circumstances, etc. This is supported by the two occurrences of ‘people are behaving’ in the BNC, and by 9 occurrences in the COCA. As for Google Books UK, out of 234 instances there was one exception, found in a book on psychology, expressing the idea that all behaviour is goal-oriented: ‘so long as people are behaving, *some* Plan or other must be executed’.

Apart from ‘people are behaving’, no unusual states of affairs attract one’s eye in the passage so much as to demand separate analysis. Films go along, people drink and talk, we watch them, directors choose. Technically, therefore, Passage 5 is reminiscent of Passage 1 in that there are no deviations that would warrant the term ‘inspired’. However, Passage 5 is sincere and Passage 1 verges on the banal. The explanation lies in the context of situation. These passages are juxtaposed in Table 9.8:

Table 9.8 Comparison of Passage 1 and Passage 5

Passage 1	Passage 5

'It seems that Martin had a reputation for sleeping with his research assistants. Most of his close colleagues must have known this, including some whom I **met socially**. Many of the people at the Memorial Service must have known. Sandra herself had been in the church that day, though she avoided meeting me afterwards, and so, she assured me, had been the girl who supplanted **her** in Martin' **affections**.'	'As you are reading, you're aware of the fact that there's only a page or two left in the book, and you get ready to close it. But with a film there's no way of telling, especially nowadays, when films are much more loosely structured, much more ambivalent, than they used to be. There's no way of telling which frame is going to be the last. The film is going along, just as life goes along, **people are behaving**, doing things, drinking, talking, and we're watching them, and at any point the director chooses, without warning, without anything being resolved, or explained, or wound up, it can just… end.

Arguably, Passage 1 is *banal* because its role in the section is to inform us of the heroine's discovery: her dead and still sorely missed husband used to have routine sex with younger women at work. Socially speaking, this is indeed banal, and the mention of others enforces the social aspects of the situation. However, at the time of such a momentous discovery one would expect the heroine to be experiencing inner conflict. She is not a reference corpus assessing how common is an affair at work. The heroine ought to be struggling between disbelief and obvious reality, between the palpable truth of twenty years of marriage and the sudden and also palpable truth of the research assistants. These inner conflicts ought to be highlighted by a reference corpus as deviations: unusual states of affairs, fractured prosodies, uncommon quasi-propositional variables, just as we saw in the work by Pushkin, Yeats, Herbert Williams (sections 5 and 7). Inner conflict always seems to be present in poetry – arguably, the majority of it is written to express and vent it. In Passage 1, Lodge does not conjure up inner conflict by language means, and therefore the passage is a failure. The preceding and following passages are different in this respect, as Lodge may have had no trouble imagining how it feels to be upset, shocked and hurt. However, he, as a writer, may not have been familiar with the experience of being hurt when finding out a shocking truth. What Lodge can convey in terms of meaning he cannot support in shades and nuances of language. The passage is ***banal***. Its function is to propel the plot onwards.

It is significant that Passage 5 is the very last one in *Changing Places*. In terms of QPVs and states of affairs it is approximately as *sincere* as Passage 1. Although two absolutely identical texts can never be found, these passages may be classified as ***sincere*** because they contain no fractured prosodies and because the QPVs of their grammatical strings are in accordance with the language norm. However, Philip's context of situation does not demand inner conflict. Philip is expressing his view on the difference between films and books – and how films are as unpredictable as life because both do not reveal signs of ending by the number of pages left until the back cover. At this point, Philip is viewed by Lodge as a film character, and Lodge the script writer decides to bring the story to an unexpected close. Philip expresses his sincere views which are remote enough from his immediate circumstances not to make his language deviate from the norm to show fluctuation of emotion. Philip the film character is unaware that Lodge is about to shout 'cut'. We are left to wonder, in the middle of Swallow's sentence, how the two couples' entangled relationships and careers will be resolved, and we might regret the ending of a good book. But Swallow as a character is fine. The passage is *sincere*.

9.4 Inspired insincerity in *Changing Places*

If by *inspired insincerity* we have come to mean subtextual deviation in quality discourse, then the following passage is a valid example:

> In Morris Zapp's view, the root of all critical error was a naïve confusion of literature with life. Life was transparent, literature opaque. Life was an open, literature a closed system. Life was composed of things, literature of words. Life was what it appeared to be about: if you were afraid your plane would crash it was about death, if you were trying to get a girl into bed it was about sex. Literature was never about what it appeared to be about, though in the case of the novel considerable ingenuity and perception were needed to crack the code of realistic illusion, which was why he had been professionally attracted to the genre (Lodge 1975: 47).

Passage 6

Milojkovic (2013a) concludes, after an extensive analysis of the strings ‘was what it *ed to be’, ‘what it * to be’ and ‘what it *ed’, that the grammatical string chosen by Morris Zapp is usually employed in cases where things are NOT what they * to be – in other words, in quite opposite situations. The QPV ‘appear’ is frequent in ‘what it *ed to be’, but not in ‘was what it *ed to be’. In fact, ‘what it appeared’ is never preceded by ‘was’ unless there is either an intervening ‘not’, or an intervening ‘exactly’ or an adverb of similar meaning. A forty-year old literary critic cannot have spent a lifetime not noticing that life is occasionally quite complex. This refusal to accept the obvious undesirable truth and sticking with bravado in the face of an impending second divorce, as well as the general feeling of impotence in a professional as well as a more personal sense, is highlighted by the semantic aura of the string.

This passage belongs to the inspired insincerity section of the cline, and is in fact on the borderline dividing poetic from non-poetic discourse. But for the tense denoting internal monologue, the passage could be uttered by any living person confused about life. The absence of fractured prosodies or unusual states of affairs contributes to this impression. If the mismatch were collocational (prosodic) and not subtextual, the (uninspired) insincerity would be inadvertent in the sense in which the word ‘symptomatic’ used by the Director of the British Council from Louw (1993: 170) is inadvertent: not willing to lie, the speaker in nevertheless ‘at pains to conceal’.

It follows from the previous sections that writing prose from the point of view of a character whose language will show marks of *inspired insincerity* demands a great gift. This kind of language will resemble the language of a persona in a poem by, shall we say, Yeats or Larkin; but, while the persona may borrow the inner conflict(s) from the writer, and indeed may vent the writer’s personal frustration or express the writer’s impressions, the novel personage will need to express frustrations that are *invented*. Perhaps research would show that such moments of *inspired (in)sincerity* mostly reflect writers’ own experiences; perhaps it is the sheer talent of the authors that will produce such gifted writing. But it is not impossible to imagine that maybe Passage 1 is less of a success than Passages 2, 3, or 4, each moving further away from banality on the inspiration cline,

because the author had never lived through the peculiar situation described in Passage 1, but described it nonetheless to satisfy the demands of the plot. Passage 6, then is an example of gifted writing that might have been inspired by a real-life prototype.

Passage 6 is an example of the 'wishful thinking' type of insincerity: hitherto Zapp has had ample proof that life is not simple, but he sticks with his pet illusion. Passage 7, which is to follow, comes from Philip's letter to Hilary and will illustrate another type of a character's insincerity, which is best described as 'uneasiness at the mention to one's wife of the woman one has had sex with':

> You'll be surprised to learn that Charles Boon is living with me at the moment! He had to leave his previous quarters at short notice due to a fire and I offered to put him up temporarily at the *request* of his *girl friend*, who lives downstairs. I can't say he's applied himself very energetically to looking for a new apartment, but he's not much trouble to me as he sleeps most of the day and is out most of the night (Lodge 1975: 133).

Passage 7

The grammatical string mentioning Melanie is the underlined 'at the * of his *'. Corpus findings in the BNC and the COCA follow:

Table 9.9 The first ten most frequent instances of 'at the * of his *' in the BNC and the COCA

	BNC	COCA
1	at the **time** of his **death** (55)	at the **time** of his **death** (179)
2	at the **top** of his **voice** (39)	at the **top** of his **lungs** (133)
3	at the **back** of his **mind** (33)	at the **end** of his **life** (103)
4	at the **end** of his **life** (28)	at the **top** of his **game** (95)
5	at the **height** of his **powers** (20)	at the **back** of his **neck** (67)
6	at the **back** of his **head** (16)	at the **top** of his **voice** (61)
7	at the **corners** of his **mouth** (16)	at the **corners** of his **mouth** (48)

8	at the **beginning** of his **reign** (12)	at the **foot** of his **bed** (46)
9	at the **beginning** of his **career** (12)	at the **back** of his **head** (42)
10	at the **nape** of his **neck** (11)	at the **top** of his **list** (38)

The first ten QPVs in the first lexical slot all denote a position, in temporal or spatial terms. The QPV 'request' does come up in the BNC and is at position 91: 'at the request of his family'. It is not among the first hundred in the COCA; however, number 77 is 'at the urging of his wife'.

It is significant that the first string to come up, 'at the time of his death', has a negative SP found in the second lexical slot. The same chunk is the most frequent in the American corpus.

The second most frequent chunk to come up is 'at the top of his lungs' in the COCA and 'at the top of his voice' in the BNC. 'At the back of his mind', the third frequent British string, remarkably relates to Philip's state of mind at the time of writing his wife a Melanie-mentioning letter; the fourth British and the third American most frequent string is 'at the end of his life', which is reminiscent of 'at the time of his death', number 1 in both.

In Google Books: One Million Books the second quasi-propositional variable cannot be recovered, as the string's length is limited to five words. These are the first twenty occurrences of 'at the * of *':

Table 9.10 The first twenty occurrences of 'at the * of his *' in the Google Books: One Million Books Corpus

1	at the **time** of his (256 868)	11	at the **hands** of his (21 055)
2	at the **head** of his (164 818)	12	at the **risk** of his (20 557)
3	at the **end** of his (123 617)	13	at the **moment** of his (19 672)
4	at the **top** of his (56 633)	14	at the **bottom** of his (18 553)
5	at the **beginning** of his (49 669	15	at the **outset** of his (17 375)
6	at the **height** of his 46759	16	at the **commencement** of his (15 868)
7	at the **close** of his 38944	17	at the **feet** of his (14 340)
8	at the **expense** of his 38794	18	at the **death** of his (14 086)
9	at the **back** of his 25648	19	at the **request** of his (13 404)
10	at the **door** of his 23125	20	at the **sight** of his (12 987)

The Google Books: One Million Books results agree with the BNC and COCA findings: most QPVs are those denoting time and space. However, there are differences, which may or may not be attributed to the fact that the second QPV could not be retrieved. The QPV 'request' appears at position 19, so 20 lines have been included in the overview instead of 10. The following QPVs break the time and space pattern :

position 8 'at the **expense** of his'
position 11 'at the **hands** of his'
position 12 'at the **risk** of his'
position 19 'at the **request** of his'
position 20 'at the **sight** of his'

The first three QPVs allude to injustice or danger. This brings us back to the negative SP of the most frequent grammatical strings in the BNC and COCA, referring to the time of death, the end of life, or even the top of voice/lungs.

Philip Swallow uses a lexical item in the key grammatical string that is not the most frequent and is to be found around position 20 in the three Google Books corpora. He may have resorted to the rather formal word 'request' in order to make it known to Hilary – not necessarily consciously, nor is the conscious nature of the statement the issue here – that the girl friend who resides downstairs is only Boon's, and that the Boon-related request is the only basis for their communication. In a random selection of contexts of 'at the request of his' in the Google Books One Million Books corpus, the second QPV mostly denotes a family member or another individual with a strong connection to the subject of the sentence. Thus, 'girl friend' fits in well, but the contexts seem to be formal rather than chatty. Hoping to create an impression of distance, Philip creates an impression of oddity by using the rare and formal QPV 'request' in a grammatical string whose aura of meaning is far from neutral and chatty: if the QPV does not specify time or place, the semantic prosody of 'at the * of his' is that of danger. The impression of oddity is bound to be

registered by a native speaker at some level, and, indeed, it was the reason for investigating the string in question.

Speaking of subtext as a pointer to the character's confusion or lack thereof, Louw (2010b) comes to mind when one reads one of Zapp's letters in *Changing Places*. In the novel, after Swallow takes in Charles Boon, both Hilary and Zapp receive poison-pen letters saying that Swallow is having an affair with Melanie. Désirée Zapp arrives to check on Zapp's daughter and finds both Swallow and Boon in the apartment. She reports the situation to Zapp, who assumes the worst and replies as follows:

> Yes, they're old buddies, so it's all too probable they've got some very corrupt scene going there with Melanie. Poor little Melanie. I feel really bad about her. I mean I didn't suppose she was still a virgin or anything, but that is no *life* for a young girl, being passed from one guy to another. Maybe if you and I could make a fresh start, Désirée, she would come and live with us (Lodge 1975: 146).

Passage 8

In his analysis of Yeats's 'Sailing to Byzantium', Louw (2010b) analyses the subtext of the first verse of the poem, 'That is no country for old men' (see also sections 2.6.2.3, Section 5.3, 5.5). He finds that the most frequent quasi-propositional variables of the string 'that is no * for' are 'reason' and 'excuse'. He sees it as the subtext of the whole poem, noting that, notwithstanding, the persona gives his own reasons and does sail for Byzantium in verse 15. 'That is no reason/excuse for such a step' is, according to Louw, the subtext of verse 1, as well as of the whole poem, and may be used to interrogate the persona – to prompt the poem's analysis.

Morris Zapp, in his reply to Désirée, uses the same grammatical string as Yeats's persona, and, as its subtext shows, in all sincerity. His words could be read as: 'Poor little Melanie. I feel really bad about her. I mean I didn't suppose she was still a virgin or anything, but that is no *reason/excuse* for a young girl, being passed from one guy to another.' While Swallow experiences inner conflict in Passage 7, Zapp knows quite well what he is talking about in Passage 8.

The inner conflict in Passage 7 Swallow describes later on, in another letter to Hilary, as 'guilty conscience':

> I'm not going to try and justify or excuse myself. I was wretched afterwards, thinking what I'd done to you […] In the circumstances, there seemed no point in upsetting you by saying anything about the episode, and it began to sink into oblivion. When I got your letter *it revived my guilty conscience*, though I didn't connect Melanie with Morris Zapp for a moment. […] But it put me in a difficult moral dilemma (Lodge 1975: 143, my italics).

Passage 9

Thus, Zapp's in Passage 6 is a 'wishful thinking' type of insincerity, and this is a 'guilty conscience' type. Note how in this case we have the opportunity of using the author's (character's) comment to arrive at a definition of fractured subtext. However, the 'guilty conscience' type in cases of adultery does seem to have some fear in it, judging by Philip's string's subtext. Could Philip be afraid of an angry outburst, a violent scene full of accusations? In Hilary's letter to Philip, the most hurting remark I could find had to do with Philip's invitation to Hilary to come and spend three weeks with him. She replies:

> But quite apart from the expense and the problem of the children, Philip, I don't think I would want to fly out anyway. I've read through your letter very carefully and I'm afraid I can't avoid the conclusion that you desire my presence mainly for the purpose of lawful sexual intercourse. I suppose you've been frightened off attempting any more extra-marital adventures, but the Euphoric spring has heated your blood to the extent that you are prepared to fly me six thousand miles to obtain relief. I'm afraid I'd find it a strain coming over in that kind of context, Philip. Even the 17-day excursion costs £ 165-15-6, <u>and nothing I can</u> do in bed could possibly be worth that money (Lodge (1975: 150).

Passage 10

The string ‘and nothing I can’ appears once in the BNC with the QPV ‘say’, and four times in the COCA with the QPV ‘do’. In the Google Books corpora, UK contains the QVP ‘do’ (65 occurrences), and the other two contain both ‘say’ (One Million Books 119, US 187) and ‘do’ (One Million Books 90, US 174). These are the only QPVs appearing in this grammatical string.

These are the first ten random contexts of ‘and nothing I can say’ (the string appears in italics in the corpus, and the relevant collocates of the string have been highlighted in bold):

1. Phillippine Short Stories 1941-1955: 1950-1955 - Page 74
books.google.com/books?isbn=9715420850
Leopoldo Y. Yabes - 2009
“Talk cannot get you out of your muddle, *and nothing I can say* **can help** you.” “Perhaps, if we talk about it a little, this muddle will clear up, and I'll be able to see my way through.” “I'd like to help you, my nephew. I mean to encourage you, but if I…

2. The Distance Between US
books.google.com/books?isbn=0758226977
Bart Yates - 2009
Guilt and worry and fear are circling around him like large, hungry buzzards, *and nothing I can say* **will change the truth of that**. I know these predatory birds all too well, and they are impervious to reason, deaf to supplication, and fully armored…

3. Conflict of Interest - Page 81
books.google.com/books?isbn=1935753320
J M Snyder - 2010
Father Nate sighed. Steepling his hands in front of him, he frowned at Alex. “It sounds like you've already made up your mind.” “I have,” Alex said, nodding. It was better this way. It was. “*And nothing I can say* **will change it,**” the priest continued.

4. A Guide To Flexible Dieting - Page 3
books.google.com/books?isbn=0967145651
Lyle McDonald - 2005

That doesn't change *and nothing I can say* **or do changes that**. What I mainly want to talk about is ways to make the reality of long-term dieting, the 'repeat forever' part a little easier to cope with psychologically. But I'm getting ahead of myself…

5. Life of Lord Kitchener, Volume 2 - Volume 2 - Page 58
books.google.com/books?isbn=1602062498
Sir George Arthur – 2007
As I said in a telegram to Brodrick, I consider this premature; already the Boers have swept away Goold- Adams's1 stock and cultivation more than once. It is merely a gift to the Boers [for us] to go too fast, *and nothing I can say* will check the…

6. The Bonds That Bind - Page 155
books.google.com/books?isbn=1462010873
Seth Giolle - 2011
"*And nothing I can say* **or do, nothing anyone can say or do, will change that.** But Eltoq and Felissa followed you of their own free will. And your intentions were valiant even if your plan and execution were flawed." Grocia felt the words land on…

7. Popular Science - Jun 1974 - Page 58
books.google.com/books?id=iFzpLpXjYdkC
Vol. 204, No. 6 - Magazine
First, your engine is too small for steady legal-speed, high-altitude driving with the load you have, *and nothing I can say* **will change that.** You can do something to help it, though. The biggest help would be to lower rear-axle ratio, about 3.90 or…

8. Never Been Kissed: A Novel - Page 170
books.google.com/books?isbn=0800732596
Melody Carlson - 2011
It's like what's been done is done *and nothing I can say* **or do—even proving my innocence in a court of law—will ever undo it.** Because of the research I did at Grandma's (as well as the lectures I received from the police and guidance…

9. Catching Fire (The Second Book of the Hunger Games) - Page 366
books.google.com/books?isbn=0439023491
Suzanne Collins – 2010

The wave should clear those away. There's nothing to do now but wait. Peeta and I sit at the edge of the water, hand in hand, wordless. He gave his speech last night but it didn't change my mind, *and nothing I can say* **will change his**. The time…

10. Speed Dialing The Dead - Page 217
books.google.com/books?isbn=059549742X
Boyd Russell - 2008

You've already made your own history *and nothing I can say* **will change that**." There was a moment of silence before Edith continued, then she began, "This has to be difficult for you, and I understand. When I watched that disk I was reviled…

Figure 9.9 The first ten random contexts of 'and nothing I can say' from the Google Books: One Million Books corpus

In seven contexts out of ten, the most frequent collocate of 'nothing' is 'change'. Two more are 'help' and 'undo'. The collocation 'say or do' features in three contexts out of the ten. Let us now look at nine random contexts of 'and nothing I can do' from the same source:

1. Honor's Bride - Page 258
books.google.com/books?isbn=0373290322
Gayle Wilson - 1998
And nothing I can do. Haviland is **determined**. And you can't in honor reveal Michael's actions which led to his death on that mission, any more than I can." He was right, she realized. Although she had sworn no oath, she, too, was bound by the ...

2. Coming full circle: the process of decolonization among post-1965 ... - Page 18
books.google.com/books?isbn=9718832432
Leny Mendoza Strobel - 2001
There is nothing I can do **about yesterday** *and nothing I can do* **about tomorrow**. Today is all I have. Yet no sooner have I declared this love for my Filipino self than I am once again criticized for being nostalgic or romantic about a past that can ...

3. Mother, mother - Page 158
books.google.com/books?isbn=0312925433

James Stingley - 1981

It's something you wouldn't have done. It's something I would never do. "You know, Mama, you and I are never going to be close like a Walt Disney family, because we just ain't. There's **nothing you can do about it**, *and nothing I can do* about it.

4. Destined to Meet - Page 82
books.google.com/books?isbn=0373032560
Jessica Steele – 1993
'I'm sorry,' she said sincerely, and feeling she carried a lot of the blame, '**Is there anything I can do? Nothing you can do?**' 'Not while Rosalind has this particular bee in her bonnet. *And nothing I can do* while the family badger me to marry but ...

5. Man on the Border - Page 24
books.google.com/books?isbn=0425195236
Dave Austin - 2004
Without some document to go on, **there is nothing he can do** *and nothing I can do*. The only thing I'd disagree with is him telling you to get a lawyer. That would be money thrown away." "But Sheriff, that will was stolen. Leon took it out of the box ...

6. An Introduction to Cognitive Behaviour Therapy: Skills *and* ...
books.google.com/books?isbn=141290840X
David Westbrook, Helen Kennerley, Joan Kirk - 2007
... unlovable, incompetent, worthless, a failure, etc.); the world is bad (nothing good happens, life is just a series of trials); and the future is also bad (not only are myself and the world bad, but it will always be like this *and nothing I can do* will make any difference ...

7. On the Wild Side - Page 138
books.google.com/books?isbn=0553445138
Linda Warren - 1995
And nothing I can do **will change . . . anything**." She got up quickly and retrieved her sweatshirt, then started back toward the track where they'd left the horses, walking quickly, then running. Bill stood up to watch her. Helplessness, anger, and ...

8. Avatar Book One - Page 106

books.google.com/books?isbn=074340050X
S.D. Perry - 2001
And nothing I can do, **nothing at all**. For all I've researched and studied, for all my abilities, I may as well wish them better: There had been six in all. Now there were three in the infirmary's ICU, the room quiet and still but for monitor sounds and the ...

9. Team of Rivals: The Political Genius of Abraham Lincoln - Page 754
books.google.com/books?isbn=1416549838
Doris Kearns Goodwin - 2005 - Preview - More editions
"You could hardly believe it possible, but my mother protests to me that she is in actual want *and nothing I can do* **or say** will convince her to the contrary." Her increasingly erratic behavior persuaded Robert to commit her to a state hospital for ...

Figure 9.10 Random contexts of 'and nothing I can do' from the Google Books: One Million Books

It will come as no surprise to find that 'nothing I can do' is an expression of helplessness, despite the despair and the anger, in a situation where action aimed at a solution would be desirable.

Philip need not fear Hilary's reaction. She is not being unreasonable. She is putting on a brave face while knowing that nothing she can say or do can change what has taken place, no matter how much she wanted or needed it. Their marriage has been harmed beyond any ability to change it back to what it was.

It is clear that there is a connection between the type of text containing deviation from the subtext of a grammatical string and the term 'inspired insincerity' assigned to it on the cline. Passage 6 (containing the key phrase 'life was what it appeared to be about'), and Passage 7 ('at the request of his girlfriend') are examples of the character hiding the truth. Morris Zapp hides it from himself (wishful thinking), while Philip Swallow is at 'pains to conceal' (Louw 1993) it from Hilary. The subtext of Passage 10 (containing 'and nothing I can do') veils Hilary's greater disappointment than she cares to express, insisting on the triviality of expenses rather than on the irreparable harm done to the marriage. Novel writing demands effort and concentration; nevertheless, the author could not have assumed

that the very grammatical strings of his choice carried additional layers of meaning. When it comes to fractures in semantic prosodies, lexical choices are less opaque and will probably be supported by other collocates or context clues in novel writing – hence the term 'inspired sincerity'. More research into these divisions is necessary, but it is clear that the subtext of grammatical strings is an important and still unrecognized vehicle of characterization.

9.5 Authorial insincerity, forced expression and manipulative language: quality vs. non-quality discourse

According to the cline that is offered at the beginning of this section, the most inspired parts of texts will tend to contain the greatest (number of) prosodic and contextual clashes. That will also make them the most 'insincere', according to the same cline, because the oddest usages will 'betray' content that the persona/character/author will expose unawares. The writing at this point will reach the highest quality and inspiredness, as the conscious and subconscious messages will intertwine so as to leave a lasting impression on the reader/hearer.

How will this sort of insincerity – unintentional, poetic, fictitious – differ from what is usually meant by insincerity, or lie? Authorial insincerity, showing clashes with the norm, will be the product of inner pressure. Unresolved inner conflicts, unrealized insight will struggle out and transform, or *de*form, standards of lexical and lexico-grammatical collocation. As there are numerous ways to be unhappy in a marriage (Tolstoy, the opening of *Anna Karenina*), there are numerous ways to transgress a language norm under the inner pressure to express unresolved content. The fractured prosody will tend to be negative, though in a million ways, as happy people do not tend to write (Louw 2000), or, arguably, to write well. The logic will reflect Tolstoy's rule of happy and unhappy marriages – just as happy people will express themselves predictably, unhappy ones will transform a predictable event into a visible or invisible device. When Lodge is in the danger of being banal, he is happy; the misfortune he describes has not touched his heart (Passage 1). This does not necessarily mean that gifted writing is based on personal experience; it means that

gifted writing is unpredictable because *inspired*. The more unpredictable, the more so. The more unpredictable, the higher the quality.

Unlike authorial insincerity, ordinary insincerity, which is equated either with forced expression or with lies and called *uninspired insincerity* on the cline at the beginning of the section, results from outer pressure – to be more precise, pressure that has to do with persuading another. The aim is not to express inner content, but to convince another of what the inner self knows to be false – in a way, to create inner content within another that will suit the purposes of the self. When taken to extremes, ordinary insincerity is manipulative, in which it is similar to cliché. Clichés are a form of banality which always arises from a need that has to do with another – whether in an uninspired pupil who needs to write an essay on a given topic, or an uninspired journalist who needs to produce an article to a deadline.

A hypothesis was offered in Louw and Milojkovic (2016: 300) that the difference between *inspired insincerity* and *uninsipired insincerity* lies in two levels of distinction:

- at the level of intention, whether the subject hides the truth from himself/herself or from others
- at the level of language, the truth s/he hides from himself/herself will be reflected in odd QPVs, but the truth he is hiding from others is likely to be reflected in collocational mismatches, or clichés, thus turning the cline into a circle.

As was shown in Section 8 (see also Милойкович in press 2019a), this assumption has not been confirmed; instead, research to date has only confirmed the hypothesis that hiding truth from others may result in using set phrases rather than miscollocations. The mismatch in such cases is not within the collocation, but of contextual nature: the set phrases used are found in the reference corpus as part of different contexts of situation from the one where they are deployed by the speaker/liar.

In view of new findings to do with manipulative discourse (Милойкович in press 2019a) discussed in Section 8, the section on the cline termed ‘uninspired insincerity’ needs a further division into forced and manipulative discourse. ***Forced discourse*** will relate to any situation in which text is produced to please another, but where inner motivation is

lacking. Such situations have been described in Louw's example of 'symptomatic of ' (1993: 170) and Hunston's (2007: 261-263) of 'to the point of'': the speaker and writer were giving praise in accordance with their professional task, and the praise may not have been quite genuine. These utterances are similar to so called 'white lies', in which manipulation is not part of the speaker's/writer's intention, though the intention to deceive is more conscious than when it comes to Freudian slips. Translators, who are necessarily under an obligation to convey a message inspired by another, may produce text that will fall into this category – this could be especially true of poetic discourse (see Section 6).

In this connection, it is worth reminding the reader that the very term *inspired* is a technical term, intended to describe a certain kind of writing. It by no means suggests that certain kinds of writing are automatically accompanied by the presence or absence of certain degrees of inspiration. When a politician uses a cliché to persuade the nation, he may well be inspired, or motivated, but he is as insincere as a person lying downright, because clichés in that context are general truths spoken in order to influence public opinion. Inspired liars, including advertising agencies and bank employees, mean to chunk believable states of affairs so as to persuade another. This is perhaps why ***manipulative discourse*** has been shown (Section 8, Милойкович in press 2019a) to use set phrases that are ripped out of a different type of context rather than produce collocational mismatches. Authorial, or inspired, insincerity thus emerges as untruthful discourse whose purpose is to achieve certain subtle readerly effects. Unlike forced or manipulative language, authorial insincerity is part of quality discourse.

9.6 The place of rhetorical devices on the 'insincerity cline'

As the cline offered in Section 9.2 purports, at least in theory, to classify all discourse according to the presence or absence of prosodic and subtextual clashes in their contextual environment, a place on it must be assigned for what is traditionally termed 'rhetorical devices'. As much research is necessary at this stage as to what constitutes a device, and the mechanisms of chunking various devices, the solution proposed here is bound to be of highly speculative nature. It also has to be based on nothing else than the

corpus-attested definitions of devices given in Section 7, which arose out of corpus investigations of six examples of pun, two of metaphor and one of personification.

Based on the definitions given in Section 7, it seems reasonable that on the cline of prosodic and subtextual clashes rhetorical devices should be placed between inspired sincerity and inspired insincerity. It is debatable to what extent a device can be deliberate (for a discussion of deliberate metaphor, see Steen 2015: 316-317), but when it comes to devices, meaning is undoubtedly advertent. Intentionality of meaning is what distinguishes, for example, inspired sincerity ('bent on self-improvement') from forced praise ('symptomatic of'), mentioned in Louw as irony vs. insincerity. As such relexicalised discourse may rely on both lexical miscollocation and/or subtextual deviation, it should occupy the area between inspired sincerity and inspired insincerity.

9.7 Conclusion

The cline in Section 9.2 is an attempt at the classification of texts according to the nature of their deviation from the reference corpora. Although the cline does clarify how texts may differ from the point of view of CPT, the actual texts will be much more nuanced as to their place on the cline. The proposed cline classifies texts containing prosodic clashes (fractured semantic prosodies) as showing *inspired sincerity*, and those containing subtextual clashes (less usual or unexpected QPVs) as showing *inspired insincerity*. Texts that show no deviation from the corpus where the context of situation seems to demand it are deemed to show *banality* – one of the texts (Passage 1) showed such remarkable lack of deviation that it could only be fully established by corpus means.

'Sincerity' and 'insincerity' are purely technical terms at this stage, as the author had no scientific proof, when proposing this terminology, that these texts are indeed more or less sincere than others. Having said that, it is worth speculating in future research that, as subtextual deviation is more difficult to perceive without recourse to corpora than a prosodic clash, the prose writer or poet may be less aware of the meaning hidden in their chosen form of expression. The term 'inspired' is also purely technical – however, experience suggests that uninspired writing is often either forced or banal. Therefore, apart from the disclaimer of any scientific connection between the cline's terminology and the

actual author-text relation, there is a possibility of bringing the two into some sort of relation in the future. So far, the following propositions seem well-founded:

- banal texts do not succeed in conjuring up the emotional change that seems to be called for in a given context of situation, and that impairs the quality of writing;

- inspired texts differ from banal ones since, even if seemingly not containing deviations, they are bound to form an original and nuanced relationship with the surrounding context clues in the context of situation.

The following section will attempt to lend practical verification to the theoretical claims made here. An actual author will be consulted about the presence of inspiration in his writing. The stretches of texts described by him as more inspired than others will then be analysed. Although this practical study is no more than an initial step, it had to be conducted to prove or disprove the general hypothesis that textual fluctuations are brought about by emotional ones, commonly termed inspiration. Poetry suggests itself as a suitable genre to begin looking for evidence.

10 Inspired writing: a case in point

10.1 Introduction

In the previous section, the terms *inspired* and *insincerity* were used as denoting certain types of text. The section did not claim that all liars are uninspired (many could be), that banal writing cannot be used for artistic purposes, or that the more inspired the writer, the better the quality of writing. It may very well have been the case that professor Lodge was more inspired when writing Passage 5 (the ending to *Changing Places* and a contemplation on the difference between a book and a film) than when writing Passage 4 (the humorous description of Morris hiring Dr O'Shea's attic), although I personally would like to think it was the other way around. Perhaps he was equally inspired when writing both, or inspired in different degrees – or in different ways, or he was in different moods, subsisting in the shoes of two different characters at two different points in the storyline. One cannot measure inspiration as one would temperature, not because it does not come in various degrees (every writer will confess to having been more or less inspired), and not because it is a personal matter (so is temperature). According to COBUILD, inspiration is 'a feeling of enthusiasm and encouragement you get from someone or something, which gives you new ideas to create something'. The famous dictionary may have fallen short of a perfect definition in this case, especially in its overuse of the well-known collocate of 'happen' (Louw 2000 mentions that 'something' collocating with 'happen' prospects negativity), but the word 'feeling' is sufficient to doom any attempt at measuring. The *Oxford English Dictionary* has done better: 'A breathing in or infusion of some idea, purpose, etc. into the mind; the suggestion, awakening, or creation of some feeling or impulse, esp. of an exalted kind.' A breathed-in impulse to create, especially of an exalted kind, will not succumb to measuring, and neither will a similar impulse with which the reader is left suffused – of the kind that has left me perfectly and unscientifically sure that Passage 4 is more inspired than Passage 5, and that Passage 1 is the least inspired. To quote Malinowski, 'All Art […] *which lives by knowledge and not by inspiration*, must finally resolve itself into scientific study, and there is no doubt that from all points of approach we

are driven towards a scientific theory of language...' (Malinowski 1923: 297), emphasis added).[41]

The aim of Section 9 was to offer a classification of texts, based on CPT, and to connect this classification to quality of writing. It introduced the term 'inspired' into CPT to facilitate the classification, in which the following convention is established: *inspired* writing will contain more prosodic and subtextual clashes than the *banal* kind. In certain contexts of situation inspired writing will be of a higher quality than banal writing. The idea to use the term *inspired*, alongside with Louw's *insincere*, came from Larkin's use of 'but when did' (see Section 4.5, see also Milojkovic 2012, 2013a) and Pushkin's *Передо мной явилась ты* (Section 5), since both authors during the process of creation expressed meanings of which they could not have been aware. This newly-formed CPT-based convention has turned a subjective term 'inspired' into an objective descriptive criterion, and nothing else, at least for the time being.

However, it would be of interest to ask a living published poet, a native speaker of the reference corpus language, to point to a section of his poem which he himself describes as inspired. As stated at the end of Section 9, this is only an initial step, given all the real-life complexities described in the Introduction (and surely there are many more yet to be encountered).

10.2 A case study: inspired writing in English

Jonathan Boulting fits the description of a living published English poet and kindly agreed to assist the pursuit of studying inspiration. The following practical study is included to provide an illustration of the CPT term *inspired* in relation to a portion of a text which its author describes as such. Re-stating that no definition of inspiration, and no

[41] Still, it is feasible to conduct a study involving poets and writers who will point out stretches of text written when they were less and more inspired. This study will then compare degrees of inspiration reported individually with corresponding texts. It might then conclude that individual writers show more deviations from the norm when they claim they were more inspired than when writing texts showing lesser deviations from the norm.

comprehensive or conclusive study of it, is attempted here, it is still interesting to see whether the CPT-based description of inspired text (unusual states of affairs and prosodic and subtextual clashes) would fit a particular author's – any author's – idea of inspiredness. My revisable hypothesis is that it would.

The poet gave me a choice of 21 recently created and still unpublished poems. The fact that they were recently created was helpful because it meant that his memory of their creation was recent too. It took me several emails to convey to the poet the idea behind my research, in the course of which communication he provided not one but three kinds of material concerning the poem of my choice: which sections he liked best, around which sections the poem started, and which sections he felt were most inspired. Below is the poem, together with the highlighted sections:

1. Sections that the poet liked most

MAY ELDERBERRY

towards noon, a creamy fleet
of flying saucers smelling of
sex and anticipation
floats low over the sermon-
surpassing stone wall which guards
the Garden of the Virgin,
She Who in white and blue ponders
an overweight Baby Jesus
horizontal in Her arms,
His helplessly enormous head.
A pilgrim on his way away
from mummification thanks
God and His lucky stars
for this vertical Lady.

2. Sections around which the lyrics started

MAY ELDERBERRY

towards noon, a creamy fleet
of flying saucers smelling of
sex and anticipation
floats low over the sermon-
surpassing stone wall which guards
the Garden of the Virgin,
She Who in white and blue ponders
an overweight Baby Jesus
horizontal in Her arms,
His helplessly enormous head.
A pilgrim on his way away
from mummification thanks
God and His lucky stars
for this vertical Lady.

3. Sections that the poet felt were most inspired

MAY ELDERBERRY

towards noon, a creamy fleet
of flying saucers smelling of
sex and anticipation
floats low over the sermon-
surpassing stone wall which guards
the Garden of the Virgin,
She Who in white and blue ponders
an overweight Baby Jesus

horizontal in Her arms,
His helplessly enormous head.
A pilgrim on his way away
from mummification thanks
God and His lucky stars
for this vertical Lady.

The material, to my mind, suggests the following. The sections which the poet liked most at the moment of our communication must be those where the form and the meaning, in his view, combined to best advantage. It is relevant that all these sections are also highlighted as most inspired. Furthermore, there are no sections which the poet liked best but which he did not consider most inspired, but not all of the sections which the author highlighted as most inspired he also liked best.

In the poem, unusual states of affairs seem to cluster in the sections highlighted as most inspired rather than not, although those considered as most inspired ones do predominate. 'Overweight' is co-selected with 'baby' as well as with 'Jesus', 'helplessly' is followed by 'enormous', 'lady' is seen as 'vertical'. In the BNC, there are 6 instances of bab* co-selected with 'overweight', and in the COCA there are two. All these contexts are concerned with pregnancies and parenting. 'Jesus' is never described as 'overweight'. 'Enormous' and 'helplessly' are co-selected only once, in the BNC, to describe a different relation of the collocates: 'she stared somewhat helplessly at the enormous four-poster bed'. 'Lady' and 'vertical' are never co-selected. As for the opening of the poem, there are 3 instances of 'fleet of flying saucers' in the COCA (none in the BNC), all in the context of extra-terrestrial beings and not flowers (although in the BNC 'flying saucers' are once compared to 'huge celestial jellyfish'). 'Fleet' is never co-selected with 'creamy'. Saucers are sometimes associated with smell (saucer* co-selected with smell* gave one context in the BNC and 2 in the COCA), but not flying ones.

One would assume that 'sermon-surpassing' ought also to have been highlighted (there are no matching records of 'sermon' co-selected with 'surpass' in either corpora), but apparently not. It is worth pointing out that the question put to the poet concerned *most*

inspired sections, so as to avoid the need to define the dividing line between inspired and uninspired text. 'Sermon-surpassing' is a happy combination of alliteration and assonance, on the one hand, and a double meaning of 'surpassing', considering the verb refers to the actual wall surrounding a place of religious worship. 'Guards the Garden' is also a happy instance of alliteration, also unhighlighted. The rest of the unhighlighted, or less inspired, section, 'floats low over the sermon-/ surpassing stone wall which guards/ the Garden of the Virgin,/ She Who in white and blue ponders', does not seem to contain any unusual combinations. 'Float*' co-selected with 'low' appear 3 times in the BNC, and 37 times in the COCA. There are 6 instances of 'floats low' in the COCA. There are 151 and 749 instances of 'stone wall' in the BNC and the COCA, respectively. Neither is the unhighlighted opening phrase, 'towards noon', at all unusual: there are 6 instances of it in the BNC, and 2 in the COCA, in the meaning in which it is used in the poem.

The only section that is highlighted first in answer to all three questions is the one which, in my view, offers the greatest opportunity for a study of inspiration. The whole poem started around it, which means that it was 'breathed in', prompting the whole lyric, spontaneously. However, it is also relevant to the lyric, as the author also highlighted it as one of the sections he liked best at the time of replying to the research question. Not all sections marked as 'the most inspired' also deserved the epithet of 'best liked'. Being truly spontaneous as well as relevant, it warrants the most attention, whatever the outcome of the corpus research.

Due to practical corpus-imposed limitations of my searches, I found I had to ask the poet in which order the parts of the image originated. This is his reply: 'As far as I can remember, first came 'the fleet of flying saucers'. You will see why when you see an elderberry in blossom. Familiarity with Graves' 'White Goddess', with its many references to magical tree lore, and which T. S. Eliot was the first to publish, also helps for understanding the poetry of many post-World War Two English (and anglophone) poets. Then 'the creamy'. Then the 'sex and anticipation' (Boulting, personal communication by email, May 2013).

'The fleet of flying saucers', therefore, was the very first image to occur, prompted by the actual sight of flowers in bloom. A BNC corpus search (Jonathan Boulting is

British) of '* of *ing *s' yielded 977 strings. The majority of these contained at least one function word in the place of a lexical slot (e.g. 'this' in the last slot), and some included a content word instead of a lexeme combined with an inflectional ending (e.g. 'thing' in the third slot). When these were removed, it turned out that out of the first 20 concordance lines 6 could be considered:

position 1 **way** of **do**ing **thing**s (80)
position 3 **ways** of **do**ing **thing**s (57)
position 10 **registration** of **fish**ing **vessel**s (17)
position 16 **institute** of **trad**ing **standard**s (11)
position 17 **number** of **work**ing **day**s (11)
position 19 **way** of **putt**ing **thing**s (11)

Figure 10.1 A BNC search of '* of *ing *s'

The most frequent QPVs in the first lexical slot are 'way' and 'ways'. The most frequent QPV in the second lexical slot is undoubtedly 'doing'. The second QVP, 'doing', does not contradict 'fishing', 'trading', 'working' and 'putting', which all express constructive action (the last one denotes constructive action involving words in all contexts).

Let us expand the search line. The BNC searchline 'a *y * of *ing *s' produced 40 strings. Seven strings where there was confusion as to the required content and functional words were excluded, as in the previous example. I also removed strings where the first lexical slot was followed by the –ly ending, instead of –y. Lastly, I removed the string 'a temporary method of servicing overseas' as 'overseas' does not consist of a lexeme and the –s ending. The final result consists of 16 concordance lines, each appearing in the corpus only once:

position 5 a **sunn**y **curve** of **breathtak**ing **view**s
position 6 a **stead**y **winner** of **golf**ing **title**s
position 7 a **stead**y **stream** of **matur**ing **loan**s
position 8 a **stead**y **series** of **disparag**ing **comment**s

position 9 a **steady flow** of **crush**ing **forehand**s

position 10 a **sorry tale** of **miss**ing **book**s

position 13 a **secondary reinforcement** of **pre-exist**ing **leaning**s

position 17 a **nasty spell** of **draw**ing **match**es

position 18 a **nasty habit** of **includ**ing **exception**s

position 19 a **nasty habit** of **devour**ing **snail**s

position 20 a **nasty habit** of **cutt**ing **corner**s

position 21 a **nasty habit** of **cover**ing **Womble**s

position 30 a **greasy set** of **play**ing **card**s

position 32 a **funny way** of **put**ting **thing**s

position 38 a **crazy set** of **juggl**ing **ball**s

position 39 a **complimentary range** of **furnish**ing **fabric**s

Figure 10.2 A BNC search of 'a *y * of *ing *s'

The most frequent QPVs are 'steady' and 'nasty'. 'Nasty' is helped in its negative aura by 'sorry', 'greasy', and 'funny' (it is negatively used in the wider context). 'Steady' is used negatively in two out of the four cases – in the other two it is used positively. 'Crazy' is also positively used: 'Prove that you can keep everything going at once with a crazy set of juggling balls – guaranteed fun after Christmas lunch! Three-ball boxed set'. Surprisingly, 'greasy' is used in a positive context of a family evening, but there is no indication that the grease on the cards is viewed with approval as such. There are no neutral adjectives here, even 'secondary' is used slightly disparagingly: 'Like the absence of useful maritime peripheries, however, they can only serve as a secondary reinforcement of pre-existing leanings to political centralism'. All in all, in the first lexical slot we find 5 positively used adjectives ('sunny', 'steady' (twice), 'crazy' and 'complimentary') and at least 10 negatively used ones ('steady' (twice), 'sorry', 'secondary', 'nasty' (five times), 'funny'). It stands to reason that, in this concordance, the negativity of the first lexeme in the string determines the negative aura of what follows, in the same way as the positive aura of the

first adjective will determine the optimistic context of the whole string. Thus, ‘sunny’ leads to ‘breathtaking’, and ‘nasty’ leads to e.g. ‘devouring’.

To gather more helpful information, I turned to the Google Books UK. As only five-word strings are allowed, I opted for ‘a *y * of *ing’. These are the first 20 lines out of 68:

position 1 a **necessar**y **condition** of **be**ing (219)

position 2 a **fanc**y **way** of **say**ing (211)

position 3 a **read**y **means** of **distinguish**ing (204)

(position 4 a very little of nothing (172))

position 5 a r**ead**y **means** of **determin**ing (164)

position 6 a **satisfactor**y **standard** of **liv**ing (153)

position 7 a **read**y **means** of **obtain**ing (149)

position 8 a **read**y **means** of **ascertain**ing (125)

position 9 a **heart**y **fit** of **laugh**ing (117)

position 10 a **funn**y **way** of **show**ing (112)

position 11 a **bod**y **capable** of **be**ing (96)

position 12 a **necessar**y **part** of being (96)

position 13 a **read**y **means** of detecting (88)

position 14 a **necessar**y **part** of learning (86)

position 15 a **satisfactor**y **way** of dealing (86)

position 16 a **contrar**y **way** of thinking (86)

position 17 a **read**y **means** of making (84)

position 18 a **happ**y **knack** of making (80)

position 19 a **necessar**y **condition** of having (76)

position 20 a **bod**y **capable** of sustaining (74)

position 21 a **read**y means of testing (74)

Figure 10.3 A search of ‘a *y * of *ing’ in the Google Books UK corpus

The constructive activity these lines are brimming with would override the negative aura visible in the previous concordance, were it not for the rightful objection that the string is one constituent shorter, which could make all the difference. A COCA concordance was attempted and did not add much information to the BNC one of ‘a *y * of *ing *s’ (Figure 10.2). Nevertheless, there is an argument in favour of taking the last concordance seriously: surely some of these uses are followed by *s in the corpus? And surely there is more than one such example per string? Let us check.

1. Cyclopedia *of* Hardy Fruits - Page 226
   books.google.com/books?isbn=1429014350
   U. Hedrick - 2008
   Young shoots of the grape offer **a ready means of distinguishing** species and varieties through their color and the amount and character of the pubescence. Shoots may be glabrous, pubescent or hairy, and even spiny. The tendril is one of the ...

2. Rabbits, Hares and Pikas: Satus Survey and Conservation Action Plan - Page 124
   books.google.com/books?isbn=2831700191
   Joseph A. Chapman, John E. C. Flux
   ... short ears (63- 109mm) provide **a ready means of distinguishing** representatives of the genus from other Southern African leporids (Robinson 1982). Distinctive, yet subtle, pelage dif- ...

3. Principles *of* the Manufacture *of* Iron and Steel: With Some Notes ... - Page 190
   books.google.com/books?isbn=110802694X
   Isaac Lowthian Bell - 2010
   ... samples (Cleveland ironstone), submitted to the action of pure carbonic oxide, had been calcined to different degrees of hardness, as **a ready means of obtaining** variations in their physical structure. The same quantities (2 grammes of each ...

4. Urban Battle Command in the 21st Century - Page 59
   books.google.com/books?isbn=0833040545
   Russell W. Glenn, Gina Kingston - 2005
   Having chaplains visit units frequently can be a significant help in this regard.42 Likewise, providing access to other lowdensity assets such as military lawyers gives lower-level commanders **a ready means of obtaining** answers to the difficult ...

5. Ideas in Chemistry and Molecular Sciences - Page 323
   books.google.com/books?isbn=352763052X
   Bruno Pignataro - 2010
   The collisional mechanism suggests, in fact, also **a ready means of detecting** proteins/enzymes that bind small molecules by simply appending the small molecule to the DNA probe, thus expanding widely, the possible E-DNA applications.

6. *A* practical treatise on brewing, and on storing *of* beer - Page 105
   books.google.com/books?id=3mazZD0pUhEC
   William Black (of Cornhill, London.) - 1835
   I therefore repeat, that an accurate smell is invaluable to a brewer, and as I have before said, a good taste is not less advantageous, since the union of the two furnishes **a ready means of detecting** errors, which might otherwise escape ...

7. Abridgment ... containing the annual message *of* the president *of* ... - Volume 1 - Page 946
   books.google.com/books?id=DulXAAAAYAAJ
   United States. Congress - 1888
   The object of these investigations, primarily, was the discovery of **a ready means of detecting** butter-substitutes, of whatever description, as manufactured or sold in violation of law ; and thin the inicroscopist claims to have successfully ...

8. University Science and Mathematics Education in Transition - Page 159
   books.google.com/books?isbn=0387098291
   Pierre R. Moret - 1980
   Learning the importance of these aspects of the culture of mathematics is part of induction into the mathematics community of practice

(Wenger, 1998) and therefore seen as **a necessary part of learning** mathematics even though strictly cultural ...

9. Reports *of* cases argued and determined in the High Court *of* ... - Volume 2 - Page 477
   books.google.com/books?id=um4DAAAAQAAJ
   Great Britain. Court of Chancery, James Russell, John Scott Eldon (Earl of) - 1830
   ... is not a separate report, yet there are some of the offices in which **a contrary way of thinking** prevails ; and the registrars consider such a certificate to be a separate report. Under these circumstances, the general principle must be followed.

10. The Boston journal *of* philosophy and the arts - Volume 1 - Page 451
    books.google.com/books?id=C2w7AQAAIAAJ
    1824
    Such a high degree of magnetic energy, says he, being obtained by a process so simple, it suggested **a ready means of making** magnets, without the use of any magnetized substance whatever, and of giving polarity to needles, so aS to ...

11. An essay on the origin and progress *of* stereotype printing: ... - Page 64
    books.google.com/books?id=v609AAAAYAAJ
    Thomas Hodgson - 1820
    It affords **a ready means of making** corrections and additions."* Rochon presented a specimen of his performance to the Academy of Sciences on the 8th of February, 1786, and an impression from it is inserted by Camus, in his Hist, de la ...

12. The Correspondence Between Sir George Gabriel Stokes and Sir ... - Volume 1 - Page 372
    books.google.com/books?isbn=0521328314
    George Gabriel Stokes, William Thomson, David B. Wilson - 1990
    He was disposed to give it a trial in the circumnavigation expedition, and thought it might be useful as **a ready means of making** soundings, whether from boats or from the ship when a mere sounding, with a small specimen of the bottom, ...

13. Anthony's photographic bulletin - Volume 28 - Page 155
books.google.com/books?id=trVIAAAAMAAJ
1897
THE UNION OF CARBON PRINTINGAND COLLOTYPE: *A READY MEANS OF MAKING* FILMS FOR THE HELIOCHROMIC TRIPLET.* By Thomas Bolas, F. C. S., F. I. C. A MONG the most forgotten of processes may be mentioned the ...

14. *A* Voyage Round the World (Illustrated Edition) - Page 87
books.google.com/books?isbn=1406829986
W H G Kingston - 2011
Cousin Silas had **a happy knack of making** friends with savages, and especially with their children. His secret, I found, was great gentleness. While Mr McRitchie, Jerry, and I sat down on a log facing the huts, he advanced slowly towards the ...

15. The New monthly magazine - Volume 13 - Page 276
books.google.com/books?id=iTkaAQAAIAAJ
1820
Prose seems easily defined; but poetry, by the questions which are still agitated about it, almost eludes definition. On the right hand and on the left I hear thr.t Pope was not a poet ; that he had **a happy knack of making** verses, but that he simply ...

16. Lives *of* the Most Eminent and Scientific Men *of* Great Britain: ... - Page 141
books.google.com/books?id=zoBlAAAAMAAJ
Samuel Astley Dunham - 1837
Jonson had, indeed, **a happy knack of making** enemies. He provoked at the same time, the lawyers and the soldiers, by the freedom of his - reflections on their respective professions. Complaint being made to the master of the revels, he would, ...

17. Syndicated television: the first forty years, 1947-1987 - Page 80
books.google.com/books?isbn=0786411988
Hal Erickson - 2001
The Bishop had **a happy knack of making** points without clubbing the viewers with piety, and when compared to some of today's humorless shouters and thumpers, Bishop Sheen's happy knack was a knack and a half. The Christopher ...

18. The Saturday review *of* politics, literature, science and art - Volume 39 - Page 515
books.google.com/books?id=bNY9AQAAIAAJ
John Douglas Cook, Philip Harwood, Walter Herries Pollock - 1875
In short, while he knows his own place and keeps it, he has **a happy knack of making** friends everywhere. Now he is poking about the fish markets at Cadiz, having stumbled on an old acquaintance among the fishermen, who insists on offering ...

19. Phrenological Magazine - Page 524
books.google.com/books?id=evEEAAAAQAAJ
1885
Sprightly, lively, genial, almost spiritual, with **a happy knack of making** friends and making herself agreeable to them; with not much wit, but with a fund of good-nature; easily made cross, and as easily gay again ; like a stream on a May-day, ...

20. BBC music magazine - Volume 15, Issues 7-13 - Page 45
books.google.com/books?id=bSE9AQAAIAAJ
British Broadcasting Corporation - 2007
It's easy nowadays to make too much of the 'alien world' aspect of Coleridge-Taylor's life. He was half-English to start with, knew the music business of his time thoroughly, and had **a happy knack of making** friends. Composer Havergal Brian ...

21. Ingrid Bergman, My Story - Page 162

books.google.com/books?isbn=0440140854
Ingrid Bergman, Alan Burgess - 1981
They hadn't seen anybody for months. They hadn't even got a war to fight, not a bullet went off in Alaska . . . and when they saw these five little planes coming in, they were so happy. Ingrid always has had **a happy knack of making** friends.

22. Reason and Morality - Page 125

books.google.com/books?isbn=0226288765
Alan Gewirth - 1980
... an agent he would have the generic rights, since in lacking D he lacks **a necessary condition of having** rights to freedom and well-being.

Hence, the fact that every agent necessarily holds that he has the generic rights does not prevent some ...

23. Protecting the Vulnerable: *A* Re-Analysis *of* Our Social ... - Page 179
books.google.com/books?isbn=0226302997
Robert E. Goodin - 1985
6) quite rightly goes on to query whether having interests is a sufficient rather than only **a necessary condition of having** rights. Curiously enough, almost all the earlier discussions seemed to suggest that it was sufficient.50 Surely it is not: as ...

24. Arguing about abortion - Page 101

books.google.com/books?id=3ooEAQAAIAAJ
Lewis M. Schwartz - 1993
His answer is that newspapers don't have desires, that having desires is **a necessary condition of having** interests, and that having interests is **a necessary condition of having** rights: 5. Newspapers don't have desires. (AP) 6. Only things that ...

Figure 10.4 The contexts of ‘a *y * of *ing *s’ in the Google Books UK corpus

These are all the examples I was able to find in the Google Books UK corpus. Although in many cases my search string was followed by a noun phrase with a plural noun, the nouns in such phrases were preceded by a determiner or an adjective, and only the above examples perfectly corresponded to ‘a *y * of *ing *s’. Below is the concordance I was able to make on the basis of these contexts:

1 a ready means of distinguishing species
2 a ready means of distinguishing representatives
3 a ready means of obtaining variations
4 a ready means of obtaining answers
5 a ready means of detecting proteins
6 a ready means of detecting errors
7 a ready means of detecting butter-substitutes
8 a necessary part of learning mathematics

9 a contrary way of thinking prevails
10 a ready means of making magnets
11 a ready means of making corrections
12 a ready means of making soundings
13 a ready means of making films
14 a happy knack of making friends
15 a happy knack of making friends
16 a happy knack of making friends
17 a happy knack of making friends
18 a happy knack of making friends
19 a happy knack of making verses
20 a happy knack of making enemies
21 a happy knack of making points
22 a necessary condition of having rights
23 a necessary condition of having rights
24 a necessary condition of having rights
25 a necessary condition of having interests

Figure 10.5 A concordance of 'a *y * of *ing *s' from the Google Books UK corpus

Clearly, the most frequent collocate of 'a' and 'of' is 'ready means' (11 occurrences), followed by 'happy knack' (8 occurrences) and 'necessary condition' (4 occurrences). 'Making' is the most frequent third QPV (12 occurrences), while followed by 'having' with its 4 occurrences, 'detecting' and 'obtaining' with 3 each, and 'distinguishing' with 2. The most frequent collocate of 'of' on the right is 'making friends' – or, to put it differently, 'friends' is the most frequent collocate of the most frequent third QPV. The most frequent collocate of 'having', which is the second frequent third QPV, is 'rights'.

Based on Google Books UK corpus, consisting of 34 billion words, the above concordance may be said to override the negative aura present in some sections of the corresponding BNC concordance of 'a *y * of *ing *s' (Figure 10.2). There, the most

frequent collocate of ‘a’ and ‘of’ was ‘nasty habit’, with 4 occurrences out of 33; apart from it, ‘nasty’ (5 occurrences) and ‘steady’ (4 occurrences) were the only QPVs (both in the first lexical slot) that occurred more than once. Still, the sample is much smaller than in the Google Books UK corpus, and the negativity is not monolithic, as the ‘sunny curve of breathtaking views’ in Figure 10.2 shows. In the concordance based on the Google Books UK corpus, there were many more QPVs occurring more than once, and with much better frequency – up to 12 out of 25 occurrences for ‘making’ in the third lexical slot, or 11 out of 25 for ‘ready means’ in the first and second slot. Therefore, unlike the BNC concordance, with its 5 positive and 11 negative contexts, the Google Books UK concordance shows tendencies of meaning much more conclusively.

What is to be our conclusion? ‘Ready means’, ‘happy knack’, ‘necessary condition’. The two most frequent chunks are ‘a happy knack of making’ (8 out of 25), followed by ‘a necessary condition of having’ (4/25). The longest are ‘a happy knack of making friends’ (5/25) and ‘a necessary condition of having rights’ (3/25). ‘Ready means of’ are reminiscent of ‘way of doing things’ and ‘ways of doing things’ – the most frequent chunks by far in our first BNC concordance, of ‘* of *ing *s’ (Figure 10.1). However, the second BNC concordance, of ‘a *y * of *ing *s’ (Figure 10.2), remains as a finding pointing to a negative semantic aura, hence the word ‘arguably’ at the beginning of the final analysis in the next paragraph.

Arguably, the reference corpora showed that the semantic aura of the grammatical string selected by Boulting in ‘a creamy fleet of flying saucers’ (elderberry in blossom over the wall of the Virgin’s garden) is that of a positive, constructive and even goal-oriented activity. This runs contrary to the subsequent description of the Garden in the rest of the poem, in which collocates like ‘stone’, ‘ponder’, ‘overweight’, ‘horizontal’ and ‘helplessly’ suggest passivity. The elderberry is a symbol of life, vitality, active force, floating ‘low’ but at the same time ‘over’ – and in that way overcoming – the passivity of religion. The same contrast is repeated towards the end of the poem, where

> A pilgrim **on his way away**
> from mummification

(and this is one of the ‘liked’ as well as most inspired sections)

thanks
God and His lucky stars
for this **vertical** Lady.

Needless to say, no animate collocate of ‘vertical’ is to be discovered in the first hundred lines yielded by the BNC.

Finally, not every poem will have a crucial meaning to the writer, such as, for instance, ‘The Circus Animal’s Desertion’ may have had for Yeats. Some poems arise from observing blossoms over a wall. The passing nature of the poet’s interest may account for a lack of very crucial clarification that is available through corpora and that can be found for the poem by Yeats (Louw and Milojkovic 2014). Nevertheless, the corpus-based analysis has helped to discover a meaning that may not have been quite obvious from an intuitive reading. Flowers are live and statues are of stone – we knew that before – but to some, as to the poem’s pilgrim, statues may seem more alive and livelier than to others. Where is the poet in all this? In the subtext of the poem’s most inspired section(s).

When it comes to the term ’inspired’, it is worth pointing out that Jonathan Boulting repeatedly – and confusingly – inquired what I meant by the word which, to him, eluded clear definition:

> There are generally said to be two types of inspiration - Mozartian and Beethovenian. Mozart wrote it all down almost straight from the air. Beethoven struggled and sweated and agonised. Shakespeare was obviously a Mozartian, but he took Hamlet through quite a few revisions, we are told. Tennyson was an obsessive Beethovenian - so was Dylan Thomas, whom the lay public imagined as an inspired oracle of drunkenly spontaneous lines. Shelley was a raver, Keats worked hard. Byron did it on hock and seltzer ('spritzer'), but his model was the finicky Pope (who, though he claimed that as a child he 'lisped in numbers for the numbers came', was an industrious polisher) […]

> I mention all this because I have found that some of my most 'inspired' phrases/lines, have come with revision. Antonio Machado was one such poet, who found more inspiration when revising [...] Often a poem can start with a deceptively banal phrase that the poet happens to find on his tongue – if he notices it – it is often so banal he doesn't... but in terms of getting a poem going, one could claim that it too is 'inspiration' (Jonathan Boulting, personal communication by email).

I still believe that it is important for the researcher to know which lines prompted the poem, as it is very likely that their very spontaneity hides the poem's subtext – authorial subtext, shall I add – as the whole poem will consist of micro-contexts, many of which will yield to scientific study. But how useful is at this point Boulting's remark that such lines are often 'banal' – non-deviant, in CPT terms! On the other hand, it seems that, as a reader, I am likelier to mistake the passage which is from my point of view the most poignant – condensed – for the most inspired, confusing it with inspiring, or inspirational. What inspires me may or may not inspire the next person – but the passage that I mark, as a reader, will contain my, reader's, hidden meanings, unresolved conflicts, subtext.

The terminology of the cline, then, will be retained, bearing in mind one important distinction. 'Inspired' in CPT will mean 'containing deviations from the norm in quality writing'. In real life, texts that do not contain deviation, because it is uncalled for in the given context of situation, may be the result of inspiration, but in CPT they will be called 'sincere'.

10.3 Conclusion

The hope was expressed in Section 9 that some interrelation could be studied between 'inspiration' as a technical term proposed by CPT in order to describe a certain type of text, and the actual feeling prompting the creation of text that is commonly termed 'inspiration'. This section has described a case study of inspired writing involving a published English poet as a first step in this direction. The obviously necessary disclaimer at this stage must reject any idea that the findings of these studies may be generalized – rather, they have been included to illustrate a possible practical approach to inspired texts in

future research. Another obvious reservation at this point is the possibility that every writer's inspiration is so individual a process – not to mention that a writer may experience different kinds of inspiration in different circumstances, moods, stages of life and so on – that no generalisable or conclusive research on inspiration may ever be conducted. To quote but one example, Jonathan Boulting distinguishes between several kinds of inspiration.

Generally speaking, inspiration and banality need not be opposites. To illustrate the point, 'I love you' may be the most inspired as well as the most clichéd phrase in all languages ever. Unlike in real life, *inspired* in the CPT sense (fraught with deviation) ought to be viewed as opposing *banal, where the context of situation demands emotional disbalance expressed through semantic and subtextual clashes*. As for the area of the cline marked 'sincerity', its difference from 'banality' is determined by the context of situation in the text, as well as – I hypothesise – by the presence or absence of inspiration in reality. After all, inspiration enables the writer to connect form to inner meaning. The product of inspiration is the truthfulness, the precision of the connection. As inner meaning is always unique, the precision of the connection will always be individual. Nevertheless, as all individuals are similarly human, it is not impossible that future research might shed light on similarities or tendencies that are characteristic of the form-meaning connection as the outcome of inspiration.

While not claiming a direct connection between the density of devices and inspiration – and not excluding a possibility that such a tendency might be discovered in the future – one still assumes that what is inspired occurred spontaneously, and therefore contains first-hand truth, or a form-meaning connection that will remain uninfluenced by second thoughts and revisions. These lines will contain the subtext of the poem from the point of view of the writer. This does not mean that objectively the whole poem's meaning is in those lines. It is everywhere, but the lines pointed up by the author as inspired will have more evidence as to personal, authorial meaning underlying the poem. This is illustrated by the analysis of Jonathan Boulting's poetry.

On the other hand, while reading the poem written by another, one may find inspiration in lines that the author would not, if asked, highlight as the most inspired. The lines that the reader finds most inspiring must, by inference, contain the combination of

form and meaning that most precisely correspond to the reader's personal truth that s/he has seen in the poem. These lines will contain the reader's subtext – which will vary from reader to reader.

The hypothesis set out in Section 9 and the actual case study forming the basis of this section suggest some reasonable conclusions on the nature of inspiration. This section has shown that

- the terminology offered in the cline may be retained for the purpose of classifying texts
- so far no obvious connection between the density or quality of fractured prosodies/subtext and the degree to which a certain section of text was more inspired than other sections has been established
- the hidden meaning revealed by reference corpora as fractured prosodies and subtext will be found in any part of the whole poem and contribute to its appreciation
- because it involves precision of the connection between one's personal truth and its expression, inspiration will escape scientific study up to a point; however, comprehensive future research, stemming from properly formulated research questions based on pre-research phases, may point to characteristics shared by all inspired texts.

11 Contextual Prosodic Theory in the stylistics classroom

11.1 Introduction

This section and the following one describe two studies conducted with students of English from the University of Belgrade. The first study, described here, was conducted with 50 students who were in their second year of English studies, in February and March 2012. The second study, described in the next section, was the sequel of the first and was conducted with 3 third-year students, its practical phase taking place in February, March and May 2013. The results of these studies were published in Louw and Milojkovic (2016), and the results of the first were additionally published in Milojkovic (2013b).

The aim of the first study was to verify Louw's principle that text reads text, and to test a methodology which had emerged from this principle. While the previous sections set out to explicate and verify this premise and methodology, these two studies describe an experiment in which the premise was presented to actual subjects, on whom the methodology was tested in a real-life situation. The first study, in effect, aimed to convey the idea of textual interpretation through absent collocates in the given context of situation, and thus mainly focused on semantic auras of both lexical items and grammar strings. The difference between the delexical and the fully lexical, and the mechanism of relexicalisation, which is explained in Section 2 and which figures prominently in literary analyses throughout the book, was not mentioned. The second study, in its methodology and choice of subjects, aimed to avoid the limitations of the first study. Its primary goal was to deal with the issue of subtext, which in the first study was attempted but abandoned. The significance of these two studies is both theoretical and practical: they do not only put to the test and re-affirm the principles of CPT and the suitability of the proposed methodology, but also describe varied real-life responses of the students. All of this may lead to the development of a more detailed classroom stylistics methodology, capitalising on the successes of the approach and avoiding the weaknesses of this first attempt.

11.2 The first classroom stylistic study

11.2.1 Aims

The original aim of the first study was to conduct a piece of classroom stylistics research implementing Louw's principles developed from 1993 to 2013. As this was the first instance of practical implementation of CPT to teaching corpus stylistics, the study attempted to answer two research questions that were interrelated. The first research question was how well-founded Louw's claim was that text reads text – that it is indeed sufficient to analyse the target text through *similar events* in the reference corpora (or in the authorial corpus if the analyst is looking for private symbolism) without much recourse to theories and without recourse to concepts. Would the students of English, unfamiliar with Louw's principles, adopt collocation as instrumentation for meaning? The second research question followed from the first, but focused on the practical side of the experiment: would the proposed methodology, based solely on collocation being instrumentation for meaning, work?

The initial hypothesis was that the methodology could be implemented in a university setting independently of whether the students were native or non-native speakers of the studied text. Even a native speaker's intuition is insufficient to establish semantic prosodies, not to mention auras of grammatical strings; however, some semantic prosodies are sometimes recognized even by non-native speakers. Thus, although non-native speakers may find more information in a reference corpus than native speakers, the nature of that information, as well as the process of discovery, is the same for both. The large reference corpus may contain a greater number of encounters with a particular instance of language use than either a native or a non-native speaker has experienced. In the case of the native speaker there may be more cases where the knowledge passes from the known, but not consciously recognized, to the recognized, but the knowledge itself is always the same. Therefore, the principle must remain the same. It is in the level of task difficulty that the adjustments must be made, because in the case of the less proficient students the gap between their personal experience in encountering a particular expression and the unedited reference corpus experience will be greater than in the case of their more proficient peers.

The specificity of Louw's theory is that it claims to be instrumentation – it insists on being equated with practice. It is possible for a theory to be sound but to require a longer explanation or more hands-on experience with the actual data. But bearing in mind Louw's insistence on instrumentation, the 'text reads text' approach was adopted without much modification and with a minimum of explanation.

11.2.2 Methodology

The study involved both quantitative and qualitative research. The quantitative phase consisted of a learning phase (5 sessions) and a testing phase (one session of approximately one hour's duration). Each session of the learning phase began with the students being given a short excerpt of text and a concordance, with a particular question to answer. After the students had written down their answers, they proceeded to discuss the text with the teacher and other students. During this dialogue, the teacher offered her interpretation based on the concordances and the target text. Still, she encouraged the expression of individual opinion, always pointing out that meaning is in the text and the concordances from the reference corpus. The 'correct' interpretation was not the teacher's. There was no preconceived correct answer that the students were supposed to seek. Their personal interpretation was the point of the session. Sometimes opinions as to the author's possible meaning differed and then a spontaneous discussion ensued. In short, the students were learning 'by doing'.

Each session of the learning phase was focused on a different type of task. Each type of task was dealt with once. The students were asked not to make any corrections to their answers during the discussion. They were also asked to sign the papers so that their progress could be monitored.

During the first session the basic corpus linguistics terminology was introduced: 'the concordance', 'the concordance line', 'the node', 'collocates', 'the 9-word window', and 'semantic prosody'. The first session only dealt with an authorial corpus (that of Philip Larkin) as that was deemed to be the easiest type of task. The next 4 sessions dealt with concordance lines taken from a reference corpus – Tim John's corpus of the *Times* newspaper for the year 1995, containing 44.5 million words. The originally intended

duration of sessions was 15 minutes, but in practice it was sometimes necessary to spend up to half an hour.

The 'testing phase' was done in one sitting, without warning. The students were given a test consisting of 5 tasks, mirroring the different type of tasks dealt with during the 'learning phase'. The estimated time of completion was 45 minutes, but the subjects were urged to work at their own pace.

The results of both the 'learning phase' and 'the testing phase' were processed at the end. A uniform marking scheme had been established for each question, according to the following criteria:

- if the analysis fulfilled the researcher's expectations, the mark was 5 (also if the analysis was different from what had been expected, but excellent and contained detailed argumentation)
- if the analysis of the concordance lines was done correctly, but no connection between the lines and the text was established, the mark was 4
- if the lines were incorrectly interpreted, e.g. the student was misled by the first line or otherwise misinterpreted the impact of some concordance lines, the mark was 3
- if the analysis was wrong altogether, the mark was 2
- if no analysis was offered the mark was zero

Attached to the test was a questionnaire consisting of 11 questions. Its aim was to get feedback on the short course the subjects had been taught. This was the qualitative part of the first study, designed to establish what views the students had formed of text-corpus interaction and of the course, whether they had liked it, how much they had understood of its purpose and methodology, whether they had found it useful and whether they would choose it if it was on offer. This opportunity to give feedback was of benefit not only to the teacher/researcher, but also to the students, who had been told at the very beginning that they would be offered a chance to express their opinions and vent out their disagreement if such was their reaction. It was understood that the methodology must have come as a surprise to all of them.

11.2.3 Background and limitations of the study

The subjects were second-year students of English, all of them in their twenties, at the English Department of the Faculty of Philology, University of Belgrade. They were two groups out of the four comprising the current generation (120 students), referred to as group B and group D, and that is how they are referred to in this study. At the end, given that not all of them attended every session, the output of 50 students was taken into account. One group had shown itself as more proficient than the other at the entrance exam. Both may be considered representative of the language level of the current generation, as the other half was also comprised of one more and one less proficient group. The study was carried out during February and March 2012, in approximately three weeks, within the framework of the 2nd year Contemporary English (Integrated Skills) course, during class time.

It was thought that the language level of the students was sufficiently high to conduct a study of concordance lines from a reference corpus for the purpose of interpreting poetic and other texts. The students' language level was thought to fluctuate between CAE (Certificate of Advanced English) and CPE (Certificate of Proficiency in English), with some students closer to Advanced, and others closer to Proficiency. None of the students had any previous knowledge of corpus linguistics, and as to stylistics, they were mainly familiar with the classifications of figures of speech, notions of register and style, and with what is generally taught in Serbian secondary schools and on undergraduate courses at Belgrade University. By the beginning of their fourth semester the students had also completed courses in phonetics (2 semesters), morphology (1 semester), and descriptive grammar (3 semesters), and had grown used to interpreting literary texts in their English Literature courses. All in all, their background, whether linguistic or literary, was deemed sufficient for the purpose of the research, while their lack of experience in corpus linguistics or modern stylistic theories was deemed an explicit advantage.

This arrangement had its faults. The students' level of proficiency in English varied from person to person, yet the grading system was uniform. Stylistics is about nuances of interpretation, and a step away from CAE towards CPE could have been crucial in this regard. The data of their knowledge available to the researcher mainly consisted of their

Integrated Skills marks gained in the previous semester(s) and her own observation as a teacher. The marks were not wholly reliable as they partly depended on students having learnt certain grammatical structures and vocabulary items (and it was theoretically possible for the most linguistically proficient student to have avoided the cramming part and scored lower marks than her less proficient but more hard-working peers). As Writing also involved extra-linguistic criteria (e.g. how well they could construct a narrative), it is the grade gained on Speaking (two minutes of continuous speech) that seems the most reliable. This grade, and the fact that group B was on the whole more proficient than group D on the entrance exam, combined with the teacher's observation, suggests that group B showed a higher level of language proficiency and was closer to, with some members even reaching, CPE.

Another limitation was time. The whole course was conducted during the Integrated Skills (Contemporary English) course and not a stylistics one, because the researcher was currently teaching that course and thus could best organize the time, provide instruction and receive feedback. Also, by that time she had established rapport with the subjects, who trusted her intentions and believed that it was their personal interpretations that she was after. Still, there was the obstacle of affective nature – not all the students had affinity for stylistic reading of literary texts, especially poetry, yet all were equally tested. If one's performance is not motivated, if a student is not interested in the outcome, can s/he be deemed a legitimate subject? Such subjects may possess all the necessary qualities and still underperform through lack of interest. The final qualitative survey was also an attempt to shed light on the issue of the students' interest.

Variations in the students' responses were another limitation. Attempts were made to standardize marking as much as possible and grades like 3.5 and even 4.8 were introduced. If an analysis exceeded the teacher's expectations in its acuteness, or included a finding which the teacher herself had overlooked, the student was given 6 points. It seemed necessary to document and take into account the difference between a correct interpretation deserving 5 points out of 5 and a remarkably insightful response.

Still, no matter how nuanced the marking was, it did not highlight some important differences. For example, if a student analysed the concordance correctly, but failed to see

the connection between it and the studied excerpt, the mark was 4. This included the cases when a student was aware that a connection could be made, but refused to make it (during discussions in group D several students insisted that a poet had a right to individual expression and not to be 'automated'). Also, in practice, the mark of 2 was given for an altogether wrong interpretation, and the mark of zero existed to be given for no answer at all – but no one was given the mark of 1. Besides, when it came to the mark of 0, it could have been earned through lack of motivation rather than ability. At least in one case, the student had approached the previous tasks so thoroughly that she gave up on the last one through sheer exhaustion. Then, the teacher's subjectivity in the presence of so many variations is always a threat to standardised marking, despite all the conscious efforts to reward similar answers similarly.

11.2.4 Implications of CPT for lexicography and classroom stylistics

Collocation is seen by Louw as a tool which constantly creates meaning through co-occurrence. Situational meaning (meaning in the context of language and culture) is created in the form of events: the target text contains an event comparable against *similar events* in the reference corpus, and therefore in the world as represented by it. A line of best fit usually subsists between the target event and those in the reference corpus. Any text can be read against the background of similar texts and the events they represent. The author's usage may be either clarified by the reference corpus, or a prosodic or subtextual clash may be observed. Alongside with other context clues, absent collocates become crucial to the interpretation of a text. These collocates are frequent in the reference corpus. If they are lexical and their absence is felt around a lexical item, we view the author's expression as a fractured semantic prosody. If they are lexical and their absence is felt within or around a grammatical string, we view it as fractured subtext. Since foreign language speakers have been less exposed to the language experience represented by the corpus, particularly to the full range of collocations in specific situations, they need more access to concordances and collocation lists to 'unpack' them than native speakers do.

During the learning phases of the research, the students were given opportunities to see these mechanisms rather than to be told that they exist, as they would have been if they

had been attending a university course. This encountered different responses. For example, semantic prosody was illustrated by a Lodge-authored passage mentioning conference-goers 'bent on self-improvement' (Louw 1993), and a concordance consisting of 10 lines. The concordance had been edited to ensure better comprehension, and presented in a larger font, with the negative collocates highlighted in bold. Every line pointed to a negative semantic prosody. However, some students remained skeptical as to the negative prosody, and were genuinely surprised when told that the prosody is actually recorded in the COBUILD dictionary. In that initial phase, they would rather believe an academic authority than a very telling piece of evidence (the edited concordance can be found in Section 2.5.1.

But not all semantic prosodies are recorded in dictionaries. That is particularly true of grammatical strings: for example, 'but then' and 'but then again' has a narrower prosody to it than explained by Cambridge (https://dictionary.cambridge.org/dictionary/english/but-then-again?q=but+then) and Longman (https://www.ldoceonline.com/dictionary/but-then-again)[42] online dictionaries. But the explanation given by a dictionary must fit every text we encounter (and a semantic prosody by definition does not extend its aura to every single case). Even if a dictionary is willing (most are) to include a mention of a tendency of use, in most cases very thorough research needs to be carried out, which will run to pages of text and justify a whole article to be devoted to it. Naturally, it may be impracticable when compiling a dictionary, although not inconceivable in the future.

Thus, not every semantic aura will be found in dictionaries, simply because the process of discovery is still ongoing, and will probably never end. The next question is what percentage of occurrences of a specific tendency of meaning will justify the expression being described as having a prosody? How far does the aura have to spread? Is 60% enough? 65%? 70%? How much is enough for a lexicographer? How much for a stylistician? What do we tell students who attempt to interpret texts? For example, Xiao and McEnery (2006: 117) state that 'give rise to' is used negatively in 46%, positively in 13%, and neutrally in 41% of instances of use. Nevertheless, the link http://dictionary.cambridge.org/dictionary/british/give-rise-to?q=give+rise+to gives the

[42] All dictionary definitions referred to in this section were checked on 13 March 2019.

following definition and example: ‘to cause something. *International support has given rise to a new optimism in the company.*' A *Longman Dictionary* entry (http://www.ldoceonline.com/dictionary/rise_2) says: '*formal* to be the reason why something happens, especially something bad or unpeasant → provoke', and gives two examles of negative use. The second edition of the *COBUILD Dictionary* chooses a less explicit definition: 'if something gives rise to an event or situation, it causes that event or situation to happen', and quotes two negative examples. As all these definitions do not really convey the state of affairs described by Xiao and McEnery (2006: 116), they are bound to give rise to misapprehension among non-native speakers – either to suggest that the use is positive (13% in reality) or negative (46% in reality). In any case, if Xiao and McEnery's data are accurate, dictionaries are not a very reliable source of reference in a corpus stylistics classroom.

It is clear that, as for corpus stylistics, the answer to this question has to do with *similar events.* The lines in the reference corpus that carry the investigated pattern *and* describe similar events are the ones to be considered. This advanced notion was not shared with the students of English. However, 10 per cent of them did mention that in a concordance certain particular lines corresponded to the line in the studied text. This must have involved their studying contextual clues in the concordance lines. This shows that at least these students could see *similar events* in the reference corpus without previous instruction.

11.2.5 Quantitative research

As already stated in Section 2.2, the main part of the experiment, quantitative research, comprised learning sessions and the test. Both took place in the classroom and did not require any preparation on the part of the students. When it came to gradual acquisition of knowledge through tasks and class discussions, the learning phase resembled regular classwork, except for the students' understanding that participation was entirely voluntary and that it was up to them either to accept or reject the propositions presented by the teacher. The test took place instead of a routine session, without warning. Both phases are described in detail in this section.

11.2.5.1 The learning phase

The learning phase consisted of 5 sessions, which took place twice a week at the beginning of every Integrated Skills lesson. The originally planned 15 minutes of each session sometimes lasted up to half an hour.

At the first session the students encountered full concordances from the authorial corpus of Philip Larkin with the nodes 'day', 'night', 'light' and 'God'.[43] Larkin's unconventional use of these words – day is viewed pessimistically, night brings relief, light appears as dark, and God is doubted – came in useful as an illustration of how meaning is created in context through co-occurrence with other collocates. The students were given basic corpus linguistics terminology – the node and the collocates as viewed in corpus linguistics, and Sinclair's 9-word window. The relevant collocates were given in bold. For example, this is the concordance with 'God' as the node:

```
MicroConcord search SW: god
80 characters per entry
Sort : 1L/SW    unshifted.
    1  that inspired it all, And made him a god. No, he would never fail. Others, of c
    2 ortraits of Sex Sun. Tree. Beginning. God in a thicket. Crown. Never-abdicated c
    3 e, musty, unignorable silence, Brewed God knows how long. Hatless, I take off My
    4 the sky, Asking to die: 'To die, dear God, before a scum of doubt Smear the whol
    5 pausing, goes into a prayer Directing God about this eye, that knee. Their heads
    6 any nights, as many dawns, If finally God grants the wish. ~2 February ~950 Dece
    7  go on before us, they Are sitting in God's house in comfort, We shall see them
    8 ey need; And famous lips interrogated God Concerning franchise in eternity; And
    9  And thought, That'll be the life; No God any more, or sweating in the dark Abou
   10 ' Let it be understood That 'somehow' God plaits up the threads, Makes 'all for
   11 , and lips bleeding. Yes, gone, thank God! Remembering each detail We toss for h
   12 tor clenched his fists And swore that God exists, Clamping his features stiff wi
   13 adio's altarlight The hurried talk to God goes on: Thy Kingdom come, Thy will be
   14 mit with his gown and dish Talking to God (who's gone too); the big wish Is to h
```

Figure 11.1 The concordance of 'god' from Larkin's authorial corpus

While the first meeting was devoted to Larkin's private symbolism, the second session introduced semantic prosody proper. The students were given the excerpt from David Lodge's novel *Small World* quoted in Louw (1993: 165). They were first asked

[43] The 4 concordances are given in full in Milojkovic (2011b): www.belgrade.bells.fil.bg.ac.rs/Bells 3.pdf

several comprehension questions, and then given chosen concordances from the *Times* corpus, in a bigger size than usual and with the significant collocates given in bold. This is the format of the task:

Name and surname:

Consider the following short passage from the novel Small World by David Lodge:

The modern conference resembles the pilgrimage of medieval Christendom in that it allows the participants to indulge themselves in all the pleasures and diversions of travel while appearing to be austerely bent on self-improvement.

1. *Explain the meaning of the passage in your own words, either in English or Serbian.*
2. *Explain the meaning of the phrase 'bent on self-improvement', either in English or Serbian*

   ... and translate it into Serbian

Look at the concordance lines of bent on taken from a large reference corpus:

1 in a society **hell** bent on achievement. Mutable thinkers don't
2 werful enchanter, bent on **bringing Arthur to his ruin** this dev
3 iding **donkeys** all bent on business, they were forcibly impress
4 cter development. Bent on change, **even to the point of shatter**
5 r world she seems bent on **conquering**. Well, I suppose that you
6 overnment is **hell** bent on demanding greater and greater protec
7 of Yoller's wood, bent on **destroying all survivors** before purs
8 he people who are bent on doing good **they can be the danger**, s
9 stic **savagery** and bent on **engulfing and drowning trapped men** a
10 **sonal safety** and bent on **escaping not only the enemy**, but the

Judging by the random lines taken from the reference corpus, how is 'bent on' usually used?

How does it influence your understanding of the line? What is the author implying?

Do you think that intuitively you felt this at the first reading?

The concordance had been edited to facilitate understanding – without the students' knowledge, as it might have discouraged them. In this session the students were taught the notion of semantic prosodies and the prosody of 'bent on' – which, they were told at the

end, is in the dictionaries. After it was elicited from the class that Lodge is in fact being ironic in this passage, Louw's irony/insincerity dichotomy was also explained to the students.

The third session focused on the notion of absent collocates. This time the relevant collocates were not given in bold and the concordances were not edited. This is the format of the task:

Name and surname _____

DRINKING SONG

(Adrian Henri)

He became more and more drunk

As the afternoon wore off.

```
MicroConcord search SW: wore off
80 characters per entry
Sort : 1R/SW    unshifted.
1 ut after a few minutes the stinging wore off and I began to enjoy the exquisi
2 y to stare at him. The novelty soon wore off, however, as Smythe persistently
3 lbeing that it was months before it wore off.   I am still trying to remember
4 have knocked him out for half a day wore off in a fraction of the time, and for
5 ing glissandi. But the novelty soon wore off. Michael Thomas disarmingly explai
6 men using implants said the effects wore off more quickly, and 29% said they ne
7 ut eventually, inevitably, the drug wore off. Some say it was Fortensky who cal
8 increase the dosage as the effects wore off.   "What we have done is to establi
9 his clean-cut approach, the novelty wore off when they realised how much pocket

MicroConcord search SW: wore on
80 characters per entry
Sort : 1R/SW    unshifted.
1 red, four years ago, and as the day wore on a repeat looked ever more probable.
2 ed Thatcher and Major. As the night wore on a swing to the right, whether or no
3  only obscured my face but, as time wore on,had a horribly isolating effect on
4  things were to change as this game wore on.   After 17 minutes, Durrant put th
5 on to the bat and off it as the day wore on.   Although he found life more diff
6  officials reported that as the day wore on an ever-growing crowd of terrified
7 Corsie's form improved as the match wore on, and to whitewash a player of Schub
8 w more perfunctory as the afternoon wore on and finally ended up with Stewart h
9  was only sustained, as the evening wore on and got colder, by the particular i
10 tive electoral history.   The night wore on and the flow of Tory setbacks moun
11  did get harder, and time certainly wore on. And on...   But first, Wimbledon.
12 d dummy runs got better as the game wore on and some of his early ones were no
13 had become even firmer as the match wore on and were not going to be pulled ar
14 ut he played ever better as the day wore on and, after Ilott had returned to d
```

COMMENT:

The students could see the role played by absent collocates, which was a new experience of just sufficient difficulty to prove interesting without being too demanding. Also, learning phrasal verbs is always of interest to non-native speakers, so the session was a success. Although the subjects had reacted with surprise to the notion of a semantic prosody during the previous encounter, this time they understood the practical advantage of establishing the meaning of a word through its collocates in the reference corpus.

The fourth session focused on grammatical strings and their prosodies. The text under study was much longer – in fact, it included a whole poem. Before they looked at the concordances, the students were first asked to judge whether the highlighted string of words was positive or negative. Only then were they given time to look at the reference corpus material, to be asked later whether the corpus excerpts had effected a change of heart. This is the format of the task (the example itself is discussed in Section 4.5):

Name and surname:

When first we met, and touching showed
How well we knew the early moves
Behind the moonlight and the frost
The excitement and the gratitude
There stood how much our meeting owed
To other meetings, other loves.

The decades of a different life
That opened past your inch-close eyes
Belonged to others, lavished, lost;
Nor could I hold you hard enough
To call my years of hunger-strife
Back for your mouth to colonise.

Admitted; and the pain is real.
But when did love not try to change ***Is this an optimistic reference to love?***
The world back to itself *– no cost,* ***YES/ NO***

No past, no people else at all –

Only what meeting made us feel,

So new, and gentle-sharp, and strange?

COMMENT (OPTIONAL):

```
MicroConcord search SW: but when did
80 characters per entry
Sort : 1R/SW    unshifted.
1  there's nothing wrong with that. But when did a car salesman ever tell you tha
2 s may make little economic sense, but when did economics really come into the e
3 yd and Rob Lowe also participate, but when did either last make a prudential ca
4 ties in both manager and country. But when did England last have success or a p
5 be the logical time to bow out.   But when did football, life, logic, Charlton
6 tiality, and has become a cliche. But when did that deter anybody?      </Group
7 abs. The rot set in after that.   But when did the present system start, and wh
8 company, making £2billion a year. But when did you last hear critics sounding o
9 hormone is, of course, a cop-out, but when did you last hear of a netball crowd
```

1 Banks, insurance brokers and estate agents sell their products and there's nothing wrong with that. **But when did a car salesman ever tell you that you would be better off walking or taking a bus?**

2 A politically imperilled Government will probably still opt to cut taxes instead. This may make little economic sense, **but when did economics really come into the equation so close to a general election?**

3 TOMMY BOY, 97 mins, PG

After Dumb and Dumber, we now have Dumbest to date. Starring Chris Farley, yet another dubious Saturday Night Live Graduate, this is not so much a comedy of errors as an error of comedy as our hero takes over the family car-brake business when his father (the much-abused Brian Dennehy) dies from over-exertion caused by marrying Bo Derek. Dan Aykroyd and Rob Lowe also participate, **but when did either last make a prudential career move?**

6 The green devotees will doubtless tune in to Witness: Beyond the Rainbow (C4, Wednesday, 9 pm), in which the daughter of a photographer killed in the sinking of the Rainbow Worrior embarks on a quest to find out more. This approach to documentary-making virtually ensures partiality, and has become a cliché. **But when did that ever deter anybody?**

QUESTIONS:

What is the tendency of meaning in sentences starting with 'but when did'?

How does it influence your understanding of Larkin's lines? Is your perception different now?

In this session, the subjects were not only given the whole concordance from the *Times* corpus, but also the wider contexts of 4 chosen concordance lines. Lines 1, 2, 3 and 6 were chosen because they were easier to understand than others, as the concordance was rather difficult, especially given the little experience with concordances on the part of the subjects.

When first asked whether they saw the highlighted expression as positive or negative, 71 per cent of the subjects first saw it as positive. Of these students, only one fourth (24,7%) changed their views completely after studying the concordance lines and the wider contexts, and concluded that the implication of the lines was in fact negative. During the discussion time the views expressed by groups B and D differed substantially. Group B claimed that the poet was intentionally ironic. Group D claimed that concordance lines in the reference corpus, especially as it was a newspaper one, had nothing whatsoever to do with the poem and the poet. The poet, they claimed, was free to use the language as he pleased. The teacher's suggestion that a grammatical string is a basic unit in a language and may therefore be studied in a newspaper corpus as well as anywhere else, failed to convince. Neither Group B nor Group D agreed with the teacher's hypothesis that the poet intended to make a positive statement while subconsciously he disbelieved in the power of love to change the world back to itself.

The fifth section of the experiment was unsuccessful. The students were introduced to the idea of subtext on the basis of John Donne's poem 'The Good Morrow'. The poem was chosen because the students were familiar with it from their literature class. From the point of view of subtext the poem is not easy to interpret. On the other hand, the choice of Yeats's 'Sailing to Byzantium', for example (Louw 2010b) would have been more appropriate from the point of view of subtext, but difficult to deal with in the classroom as the students had not yet studied it in their literature class. It was also my impression that the students may have received too much new information in a very short time. The time was limited and I did not pursue the notion of subtext further.

11.2.5.2 The testing phase

The final test consisting of 5 questions mirrored all the task types in the first 4 'learning' sessions. The first question was related to an unedited concordance with the node ‘hope’ from the authorial corpus of Philip Larkin. The second task had to do with the negative semantic prosody of the phrasal verb ‘cook up’, bearing in mind that the prosody could not be discovered on the basis of the context clues. The text itself came from Leo Jones’ *New Progress to Proficiency* (Jones 2001: 112) and had recently been studied in the classroom, so the subjects were familiar with the wider context as well, although that knowledge was not strictly necessary. The third and the fourth tasks focused on the grammatical strings ‘but then’ and ‘but what is’ respectively. In the third task the students were invited to analyse wider contexts from the *Times* corpus, which yielded a clearer picture than a regular concordance would. In the fourth the grammatical string needed to be interpreted on the basis of a concordance.[44] The fifth task was based on the third step in the learning phase, probing into the absent collocates of a phrasal verb. In this case, in order to save time and also to vary the tasks, the concordance itself was skipped. The students were given the result of the concordance analysis and asked to connect it to the studied excerpt from yet another poem by Larkin. This is the layout of the final test:

Name and last name

1. *What is Philip Larkin’s view of hope? You have before you the contexts in which he used the word ‘hope’. In his authorial corpus there exist 9 lines overall. Lines 6 and 9, from a birthday poem to a friend, and lines 1 and 4, from a jocular last will and testament, have been omitted as they belong to occasional poetry, and therefore not likely to express the poet’s true attitude.*

```
MicroConcord search SW: hope
80 characters per entry
Sort : 1R/SW    unshifted.
2  signalled in attics and gardens like Hope, And ever would pass From address to
3 claims The end of choice, the last of hope; and all Here to confess that someth
5 what I desired - That long and sickly hope, someday to be As she is - gave a fl
```

[44] Both concordances in the final test were edited – some unmotivating lines were removed – but the necessary level of difficulty was preserved.

```
7 e Through doubt from endless love and hope To hate and terror; Each in their do
8  it's a different country. All we can hope to leave them now is money. 10 lanua
```

COMMENT:

2. *Read the following familiar text from New Progress to Proficiency by Leo Jones. Then read the given concordance lines from the TIMES reference corpus.*

   'The idea of preserving biological diversity gives most people a warm feeling inside. But what, exactly, is diversity? And which kind is most worth preserving? It may be anathema to save-the-lot environmentalists who hate setting such priorities, but academics are starting to ***cook up answers****.'*

```
MicroConcord search SW: cooked up
1  and Pacific supermarket, and so I cooked up a story called A & P. I drove my d
2 h of Euro-scepticism, and you have cooked up a crisis."   Tory Euro-sceptics wi
3  a stream of mixed notices, having cooked up a storm in America. "Crime in exce
4 ister for the energy industry, had cooked up a £1.2 billion payout to them from
6  fact that this whole exercise was cooked up by a record company executive, and
7 tarting to resemble a cynical ploy cooked up by lenders to force the government
9 st demand for tax-planning schemes cooked up by Jenkins and his colleagues, whi
11 liance claimed the affair had been cooked up by the Russians in an attempt to
13 e fallen for every publicity stunt cooked up by the lawyers in the Simpson cas
17 ese than the Mayan extravaganza he cooked up. Certainly the claim that the bui
18 mmary of the predicament Slovo has cooked up for her headstrong part-time dete
20  has proved controversial, but was cooked up in close consultation with Major.
22 one knew that. Whatever scheme was cooked up, London would rally to the common
23 mmon murderer. Salvatore Cammarano cooked up the plot and later provided somet
25 said: "It was the father who cooked up the plot to say the car was stolen
27  was it. Yet the officials who had cooked up this crass plan bounced councillo
29 ant turns out to be less than it's cooked up to be, and Connie's disillusionme
30 ed up the National Lottery (I said cooked up, you sniggering lot) and departed
```

COMMENT:

3. *Read the following poem by Philip Larkin.*

   REASONS FOR ATTENDANCE

   The trumpet's voice, loud and authoritative,

   Draws me a moment to the lighted glass

   To watch the dancers - all under twenty-five -

   Solemnly on the beat of happiness.

- Or so I fancy, sensing the smoke and sweat,
The wonderful feel of girls. ***Why be out here?***
But then, why be in there? Sex, yes, but what
Is sex ? *Surely to think the lion's share*
Of happiness is found by couples – sheer

Inaccuracy, as far as I'm concerned.
What calls me is that lifted, rough-tongued bell
(Art, if you like) whose individual sound
Insists I too am individual.
It speaks; I hear; others may hear as well,

But not for me, nor I for them; and so
With happiness. Therefore I stay outside,
Believing this, and they maul to and fro,
Believing that; and both are satisfied,
If no one has misjudged himself. Or lied.

The following contexts are taken from the 1995 Times corpus:

1 It's just that art students, and art critics for that matter, spend a lot of time in galleries thinking about sex. **But then**, everyone used to go to galleries to think about sex.

2 Ibsen himself was subject to fits of depression, so he wasn't one for light entertainment. **But then**, few Norwegian entertainers are.

3 'That reminds me,' he said, 'did you translate the poem?'
I brought out a grubby piece of paper, made soft by much handling, [and read my translation].
'It's not bad,' said Daniel, 'but you didn't do the rhymes.'
'Are you kidding? Look at the rhyme scheme: a,b,b,a,b,b,c,d,c,d. It's impossible.'
Daniel sniffed. 'Paul-Jean Toulet did it,' he said. '**But then**, French is a richer language than English.'

4 The end of the Mozart story is tragic and you may even weep, as I did, as you read this affectionate account of his last days. Mozart's life could easily have been so much happier. **But then**, considering those 626 works in the Kochel catalogue (= a complete, chronological list of Mozart's works), would we really have things otherwise?

5 He found her beautiful and alluring. **But then**, eligible man-about-town Hewitt finds many women beautiful and alluring. Nobody has denied, however, that it was Diana who started the serious flirtation that led him to her bedchamber.

6 'Our love keeps us going.' It's not easy living in a Frankfurt jail, **but then again** it wasn't easy living with the guilt and angst of running the error account on Barings Bank and meddling with millions, other people's millions, as if they were Mars bars.
'Sure, I might meet someone nice, **but then again** I might meet someone I don't *want to meet.'*

Take into account the contexts you have just read. How do you understand the 'but then' line in Larkin? Give reasons based on the Times contexts.

4. *Now read the contexts of '**but what is'** found in the 1995 Times corpus:*

```
MicroConcord search SW: but what is
80 characters per entry
Sort : 1R/SW    unshifted.
1 f of 1% of total public spending. But what is a majority taste? Nothing, really
2 been there the old money of Eton, but what is a school to do with a boy who, re
3 r desk pontificating arrogantly". But what is a columnist for if not to pontifi
4 onOfPaper>   <Story>     <Group>  But what is a beer without a hangover?   Wine
36 th the price war with Wordsworth. But what is going on here? If selling 99p bo
42 are ball into the six-yard box.   But what is he meant to do, other than act a
43 preference holders. Clear enough; but what is he doing upping his stake in ano
56 aling, though possibly necessary. But what is it all for?   "Have some knowled
58 ibition spaces in central London, but what is it beyond that? The Academicians
73 r remains far removed from normal but what is normal behaviour for a king? Sei
79 oes not look like a comic genius, but what is one supposed to look like? As he
80 enstern, all we get is incidents; but what is our role in these incidents? Ha
95  claims are "exorbitant demands". But what is reasonable?   Sybil Gooldrich, o
97 rought down to earth with a bump. But what is risk, and how can you avoid it?
112  deaf members of the audience."  But what is the point of interpreting opera
114 ood as Claridge's or at my house, but what is?   The clientele were an odd mi
116 ould love to lead the Government, but what is the point if the party is too a
117 he accused sold were not genuine. But what is the difference between a genuin
118 ng as many sights as time allows. But what is the rush? Rome was not built in
119 ow inflation and a trade surplus. But what is the point? This is a question w
127 s is impressive and he is excited but what is the reality?   Market forces in
130  dangerous Rollerblades can be.   But what is the attraction? Unlike traditio
139 ically/she's using him fiscally." But what is this thing called friendship? W
```

144 ays before an international game. But what is to stop the clubs refusing to s

After you have read the concordance lines, how do you understand the 'but what is' line in Larkin? Give your reasons, basing them on the concordance lines.

5. *And everywhere the stifling mass of night*
   Swamps the bright nervous day and puts it out.

These lines come from the poem 'Midsummer Night' (Philip Larkin again) which deals with transition between day and night. The phrase 'it out' was searched in the Times corpus. In it 195 lines were found. 'Out' was mostly a particle belonging to a phrasal verb, with 'it' as its direct object, like 'carry it out', 'pull it out', 'sort it out'. Mostly the underlying argument in the concordance lines was that the action described by the phrasal verb was intended to solve a problem. 4 concordance lines contained 'put it out'. In all the 4 lines what needed to be put out was a great fire.

How would you apply this knowledge to the interpretation of the lines from the poem?

11.2.5.3 Discussion

The tests filled in during both the learning and the testing phase were marked. In the tables below, the tasks in the learning phase are marked as 'a': for example, 2a, 3a etc. They are juxtaposed with the results from the testing phase, marked as 1b, 2b etc. The column '1a' is empty because the subjects' answers were not graded during the first session. The column 4a is also empty, as task 4b was focused on a grammatical string as was 3b, therefore both 3b and 4b in the testing phase corresponded to 3a in the learning phase.

Results of both groups were entered into separate tables. The students were given marks from 0 to 5. In practice, no one was given the mark of 1. As previously stated, in cases of exceptionally astute judgement, the mark was 6 out of 5. Average results per each task were calculated, for the two groups separately as well as for all the subjects together. Average marks were calculated per each group and for all the subjects, in points as well as per cent. The final marks were plotted in a diagram, for the two groups separately as well as for all the students together.

Table 11.1 The results scored by Group B in the testing phase of the classroom stylistics research

Group B	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	sum	%
Average B		5	5	4.2	4	4		4.3	4.6	4.1	21.38	85.509
Standard deviation B		0.7	0	0.7	1	1		1	0.4	1.1	2.202	8.8095
Count B		22	9	22	20	22		22	18	22	22	22

Table 11.2 The results scored by Group D in the testing phase of the classroom stylistics research

Group D	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	sum	%
Average D		4.5	5	4.1	3	3		4.1	4	3.7	19.68	78.714
Standard deviation D		1.1	0	0.8	1	1		1.4	1	1.8	3.246	12.982
Count D		28	10	28	24	28		28	23	28	28	28

The difference in the marks of the two groups for the ‘b’ tasks is statistically significant (p=0.041). Clearly, the subjects from group D scored fewer points than those from group B. Here is the plotted curve of the distribution of the marks for both groups and for all the students together:

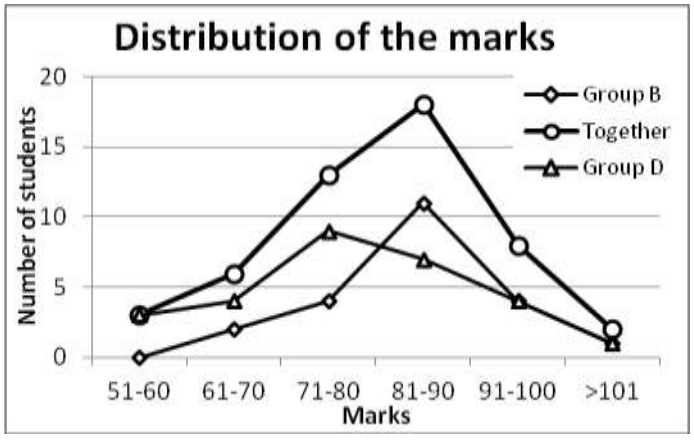

Figure 11.2 The distribution of final marks in the testing phase of the classroom stylistics research

The results scored in group D were lower than in group B, which confirmed the initial assumption that the more proficient group would score better results. Still, the results of both groups suggest that the difficulty of the tasks was adequate, as all the three curves could have been obtained after any undergraduate course of moderate difficulty.

Since the first research question was to see whether text read text for Belgrade students of English, it was important to establish the percentage of students who had scored the highest marks on the final test (5 or 6). Given the study's limitations, the figures seem to suggest that the research question was answered positively:

- Question 1b (HOPE) 62%
- Question 2b (COOKED UP) 34%
- Question 3b (BUT THEN) 14%
- Question 4b (BUT WHAT IS) 30%
- Question 5b (PUT IT OUT) 32%

The percentages of students who scored the mark of 5 or 6 are even more important for the first research question than the previously shown tables and plots, as the mark of 4 was given to students who correctly interpreted the concordances but could not or would not see the connection between the concordances and the studied text.

Another finding relevant to the first research question was that 5 students out of the 50 (10%) had made comments when studying concordance lines in 4b ('but what is') which could be construed as attempts to look for *similar events* in the reference corpus. As mentioned in Section 11.2.4, these students pointed out that certain lines in the reference corpus concordance corresponded to the line studied in the text, and it was clear from their coments that the correspondence they were referring to was based on the similarity of the two situations. As such a method of interpretation had not been mentioned in the classroom, this also confirms Louw's stance that 'text reads text'.

11.2.6 Qualitative research

The quantitative phase was followed by a questionnaire comprising 11 questions. Its aim was to see how well the students understood the point of the course, whether they had found it useful, what they thought of the methodology, whether they had enjoyed it and whether they would choose it if it were on offer as part of the curriculum:[45]

Questionnaire

Please read all the questions first before answering.

1. What is corpus linguistics?

2. What is corpus stylistics?

3. What is stylistics?

4. Would you have appreciated being given more terminology and background when doing classroom corpus stylistics?

5. Do you feel you have learnt something from this course? What?

6. What can a foreign student at your level of knowledge learn from this course?

7. In your view, what can a native speaker learn from this course?

8. What was your overall view of the teaching methodology?

9. Do you feel you have been encouraged to develop your own opinion?

10. How difficult did you find the course? What might have caused this?

11. Did you enjoy the course? If corpus stylistics was on offer at this department, would you consider choosing this subject?

The aim of the first three questions was to discover what definitions of corpus linguistics, corpus stylistics and stylistics the students would give after being exposed to the short course that in itself was based on the principle that 'text reads text'. The researcher was curious to see how the subjects understood the disciplines of corpus linguistics, corpus stylistics and stylistics. Any answer that was not wrong was marked with a 'yes', wrong answers were marked with a 'no' and the absence of the answer was marked with a '0'.

[45] The questionnaire was in part inspired by Burke (2004).

Out of the 50 students:

- 52% gave acceptable definitions of corpus linguistics,
- 48% defined corpus stylistics (6% more defined it by means of the word ‘style’)
- 32% defined stylistics (16% more defined it by means of the word ‘style’).

Answers dependent on the word ‘style’ were studied separately, as this definition seems too vague in the circumstances of this particular research, so it is not certain what the subjects actually meant and how they defined style as such.

In the subsequent questions, positive answers were marked with a ‘yes’, and negative with a ‘no’. To preserve this principle (describing answers as ‘yes’ if the feedback was positive and ‘no’ if negative), in Question 8 ‘yes*’ means that the course was not found difficult by the student, and ‘no*’ means that it was found difficult. Here follows a short analysis of the students' feedback:

- Half of the subjects (50%) suggested they would have liked more terminology and background, and 30% said they would not have wanted more.
- To all three subsequent questions (5, 6 and 7) as many as 70% of the subjects replied in the affirmative while showing sufficient understanding of the point of the course (positive answers that showed that the student did not understand the main point of the present corpus stylistics course were not marked with a ‘yes’).
- The adopted teaching methodology was approved of by 72% of the subjects.
- 48% stated that the course was not difficult.

In the last question consisting in fact of two, one related to enjoyment of the course and the other to whether the student would choose it if it was on offer, the adopted description of the answer was e.g. yes/no – the student enjoyed it but would not choose it. This is the distribution of answers:

Table 11.3 The students' responses to question 11 in the qualitative phase of the classroom stylistics research

yes/yes	yes/maybe	yes/no	no/no	no/yes	no answer
36%	16%	22%	14%	4%	8%

The results of the qualitative survey suggest that the impact of the course was overall significant, and the subjects' reaction was positive. More than 70% of the subjects claimed it was useful, approved of the methodology, and stated that they had enjoyed it. These percentages might have been higher, had the subjects' affinities and interests been consulted. All this has a bearing on the second research question asked in this study – whether the CPT-based methodology proposed in this section would prove to be successful.

11.3 Conclusion

The percentage of the students who completed the final test successfully, together with the fact that there was almost no theoretical instruction, proves that text does read text for the non-native students of English at the Belgrade English Department so far as it can reasonably be expected. Viewed against the background of the feedback provided by the questionnaire, this finding also confirmed that the second research question – whether a CPT-based methodology would prove to be successful – was answered positively. The fact that the more proficient group achieved noticeably better results suggests that native speakers would have been even more proficient, but that the principle is the same regardless of the level of proficiency. Both quantitative and qualitative results in fact exceeded the researcher's expectations, given the study's limitations, which deserve some dwelling upon. Here follows a list of the factors which may have negatively influenced the study's outcome.

- **Motivation**. As many as 36% of the students stated they would not choose this course if it were on offer. This suggests that the final results may have been even more encouraging had only the motivated students been tested. As

things stood, a fair number of students were not particularly interested in poetry, corpora or stylistics.

- **Personal maturity**. This will certainly be reflected in stylistic interpretation – both of the poem and of the concordance lines. The subjects, being in their twenties, may not have had the experience that the author tried to convey, and therefore could not interpret what they had not understood.

- **Checking understanding**. When interpreting a poem or part of it through concordance lines, it is first necessary to check each student's basic understanding and the degree of their appreciation of the text as readers. It was not done, and all students were tested in the same way. Thus, the findings may shed light on how an average generation of second year students may react to this sort of course, but for finer nuances of the process of interpretation a more detailed study ought to have been conducted.

- **Lack of theoretical background or practical skills**. No proper course of corpus stylistics would have been founded on such a minimum of instruction. This was a necessary condition for answering the first research question. However, in real life, more students might have responded to this kind of teaching positively after reading on semantic prosody. The outlook on text presented to them differed dramatically from the manner of their dealing with text and meaning throughout their previous schooling. On the other hand, the research questions demanded that the subjects should not feel under pressure to get good grades, which would have been the case had this been a proper university course.

  The lack of practical experience may have influenced the subjects' expectations. The final test was slightly more difficult because the answers were slightly less obvious. The students may have expected prosodic clashes where there were none and may not have been prepared for the other option

– that interpretation may be deepened when not changed by the concordance, especially in the case of non-native speakers of English who do not have the native speakers' accumulated experience. However, many students commented that after reading the concordance lines they understood the target line in the text better. Ironically, in some cases they claimed it even if objectively they had misunderstood the line.

A lack of basic skills in reading concordances on the part of the subjects sometimes led to overgeneralisation, when the first concordance line influenced the interpretation of the whole concordance. Another interesting error was misinterpretation based on the subject's personal experience of life. Also, the subjects of the study lacked experience in making sense of the syntactic structure of a concordance line (this difficulty is referred to in Stewart 2010: 105). All these errors would have been addressed on a proper stylistics course.

- **Reasoning behind the subjects' performance**. The marks given should be interpreted as comments on the connection perceived between the reference corpus and the authorial text under investigation. A difference must be made between a student who is not capable of perceiving how a concordance can assist interpretation, and a student who can see how it can be done, but refuses to accept that a reference corpus can be allowed to read text. Both students were given the mark of 4 and the distinction is not reflected in the results. Besides, from their comments it was sometimes difficult to see why exactly the connection between the concordance and the target text had not been made – whether the student was not capable of making the connection or refused to make it.

In retrospect, the final test may have been too difficult for a fair number of students, due not exactly to the lack of linguistic proficiency, but due to a combination of not enough English to understand all of the text and not enough general critical skills to interpret the English that they understood. It would have been sensible either to give them a poem that

had already been interpreted in their English literature class (it would have had to be one by John Donne, for example), or to go through the poem with them first to ensure comprehension (which would have been difficult because of the lack of time). The results are therefore a mixed picture of enough or not enough comprehension, enough or not enough critical skills and enough or not enough of corpus stylistics performance. This is, after all, how it would have turned out in real life, but one wishes for more concrete findings that would have taken more time than originally planned. Nonetheless, the results of the quantitative part show that the tasks were not too difficult, and those of the qualitative survey suggest that the course was appreciated by the majority of its participants.

12 Student-centred stylistics: does subtext read text?

12.1 Introduction

This section describes a case study which took place after the quantitative and qualitative research presented in Section 11. This follow-up was conducted during the next school year in order to carry classroom research further into the domain of subtext. As stated in the previous section, although subtext had initially been planned as the final part of the study presented there, that intention had to be abandoned. It was clear that both the novelty and the subtlety of subtext demanded much better focus and more dedication than the limitations of the first study could allow. Therefore, a second study was devised in order to observe the students' reception of it under more appropriate conditions.

Why did subtext prove too difficult at the first attempt? Clearly, during the 30 minutes allotted to the introduction of subtext, the students could not see the connection between the most frequent lexical variables of Donne's chosen grammar string and the context of situation in the studied line. This was thought to be the consequence of the general limitations of the first study, such as too little time spent on classroom instruction and no access to theoretical works or even other examples. However, these limitations had not stood in the way during the 4 preceding sessions, when semantic prosody, absent collocates and semantic auras of grammar strings were discussed. Besides, the poem by Donne was familiar to the students, who had recently studied it in their literature class and knew everything about its context of situation, as well as about the life of Donne himself and the characteristics of his earlier and later verse.

The answer to the question why subtext resisted comprehension must lie in its maximal intuitive opacity. Semantic prosody is less opaque, which has been referred to as twenty-twenty hindsight: 'the tendency to claim that one 'felt' the presence of a form which was inaccessible to one's intuition until it was revealed through research (Louw 1993: 173)'. The same can be said of the auras of those grammatical strings that were discussed in class (e.g. the strings 'but when did' and 'but then').

The other reason must be the slightly different approach one adopts to subtextual reading, as opposed to reading through absent but frequent collocates generally. When discussing semantic auras, one focuses on the additional nuances of meaning that are

discovered through the frequent collocates and contexts of situation in the corpus. As regards subtext, the mechanism is that of substitution rather than addition. The stylistician substitutes the given lexis in the grammar string with the frequent QPVs in the corpus and reads the poem again. For example, 'That is no country for old men' becomes 'that is no reason' or 'that is no excuse' (Louw 2010b). Arguably, the students would most probably have accepted the mechanism of substitution by absent collocates only in cases of deliberate wordplay (e.g. 'the afternoon wore off' as opposed to 'the afternoon wore on', see Section 11.2.5.1). The abstractness of the frequent QPVs could be the reason why, chronologically, this is (bar prospection) the most recent development in CPT, and why it was first arrived at through philosophy of language (Louw 2010a, 2010b) rather than intuitively.

The second study, therefore, attempted to create better conditions in which 3rd year students of English who had participated in the first study could analyse certain authorial grammatical strings against the background of reference corpora. The research questions were the same: whether the 'text reads text' principle worked for the students, and whether the proposed CPT-based methodology would prove successful. To this end, the interaction between the subjects and the researcher this time took place outside the classroom, hence the epithet 'student-centred' replacing 'classroom' in the title. This section describes the subjects' progress in mastering subtextual reading under improved conditions, recounts their own views on it, and outlines possible directions for future research.

12.2 Aims

As explained in the Introduction, the second study attempted to continue the research into whether text reads text for the Belgrade students of English, this time against the background of subtext-related tasks. The aim of the second study was to look at the students' reception of subtext without the restrictions of the first: they were to have read three relevant theoretical works by Louw and to be given enough time to process the assignments. Given that the whole process was to be more time-consuming and to occur outside the classroom, the subjects were also to be pre-selected according to ability and interest (a safeguard against another limitation mentioned in Section 11).

In particular, in the second study, the selected several students were to focus on some grammatical strings in Yeats's poem 'The Circus Animals' Desertion'. They were to use the corpus material provided by myself, which I had used to analyse the poem in Louw and Milojkovic (2014), and which is presented in Section 7. This would give me a chance to view their findings against my own existing analysis. Afterwards, the students were to read my analysis from Louw and Milojkovic (2014) in order to give feedback on the difference between their findings and mine. This feedback was meant to be of benefit both to the students, who were expected to learn more about the connection between the poem and the relevant contexts from the reference corpora, and to the researcher, who would learn more about the students' perception of subtext.

12.3 Methodology

After the first study, I approached only those students who had scored at least the mark of 5 on all the 5 tasks of the final test. Initially, all the 9 students I approached enthusiastically agreed to participate, but as the second semester of the third year progressed, 6 students decided to postpone or abandon participation, emphasising that the reason was the regular workload, which was objectively considerable. In the first phase of the second project, the students spent about 2 months reading Louw's work (1993, 2000 and 2010b) and discussing it on email. The students would read the paper and, when ready, offer their comments on it using the 'reply all' option. I participated in the discussion, but, when it came to subtext, I reduced explanation to the absolute minimum. This phase roughly corresponded to the learning phase of the first study, with the difference that no actual training related to extracting subtext took place – the subjects learnt about subtext by reading theory and examples, unlike in the first study, when the learning phase involved almost no theoretical knowledge but focused on working with concordances.

In this study, before the actual concordances were sent to the students, I asked them to read 'The Circus Animals' Desertion' and to offer any comment they wished. I was not against their reading existing critical comments on the poem in order to grasp its nuances better. In order to provide the necessary background, I also asked them to read on the life of Yeats. I did not think it wise to direct them to particular phases or events in the life of Yeats

that were relevant to the poem, because I thought it would bias their corpus stylistic interpretation.

What corresponded to the testing phase in the first study this time consisted of the subjects' commenting on several grammatical strings in Yeats's poem. When given the concordances, the students were warned not to choose the 'reply all' option this time, but to send their responses directly to me. Several times, when it seemed necessary, I prompted a student in order to see if more interpretation could be extracted, but avoided concrete prompts. The concordances I sent to the students focused on 3 grammatical strings in Yeats' poem, and were those I myself had used when working on Louw and Milojkovic (2014).

The last phase of the study involved the students' giving feedback, instead of handing in interpretation to be evaluated. After their responses to all tasks had been received, I sent them my analysis as described in my chapter, and asked them to comment on my findings, and in particular on the difference between their discoveries and mine. Thus, the last phase only roughly corresponded to the qualitative research of the first study, because it focused on the differences in textual interpretation. The whole study, with interruptions, took place roughly from November 2012 to May 2013.

12.4 Background to the study and its limitations

The students who engaged in this second study had the following advantages over the subjects of the first study:

- They were one year older, which must have influenced the level of their proficiency (one year at university may make a huge difference).
- They had already participated in a similar study the previous year.
- They truly enjoyed the experience of text interpretation through concordances from reference corpora.
- They were stylistically gifted and interested in stylistics.
- They had read three papers by Louw, bearing on semantic prosody, Contextual Prosodic Theory and subtext.

- They worked on CPT at their own convenience.
- They were additionally stimulated by having been found gifted and by helping the development of a new field of linguistics.

These were the answers to the limitations of the first study; however, the second study was not without limitations of its own. The main one of these was having to conduct a one-time piece of research with only 3 subjects. More students would have greatly enhanced the study's validity, but it is questionable if this could have been possible while retaining the above list of advantages. It was thought better to work with what resources were available and to make use of what could be found out, rather than to lose the advantage of working with gifted and motivated subjects, even if too few. Moreover, the low number of students could be turned to advantage, as it offered an opportunity of detailed processing of their responses.

Other restrictions were practical in nature:

- Sometimes too much time passed in between emails.
- Although the papers were given and the responses received, the students' understanding might have been more properly checked.
- A textbook addressing the needs of undergraduates would have been more suitable than original papers by Louw.
- The complete absence of previous practice (in other words, no 'learning phase') may have influenced final results.
- The students were not given access to as wide contexts as I had had access to in my Yeats analysis (I could always study any context in any depth).
- The students had been asked to get acquainted with the life of Yeats and to read the text of the poem. This had to be enough for text interpretation. However, I had had access to more information than they when I was writing Louw and Milojkovic (2014), and this knowledge of the macrocontext must have ensured a qualitative difference in my interpretation. The testing phase of the research is therefore less valid than it could have been, had the students availed themselves

of more facts and more critics' ideas. Indeed, their quality of interpretation may even have exceeded their teacher's.

- It was vital that the demands of the research did not obstruct the subjects' overall academic performance at university, which bears upon some of the limitations stated above.

On the whole, although sufficient insights were to be yielded from this study in its present form, more reliable and useful information could have been gained in more favourable circumstances. In particular, I would like to dwell further on the fact that I was so keen on finding out what the students' genuine reaction to subtext would be that I omitted any practical tutorship before the actual testing phase. I had only checked that the subjects' comprehension of the phenomenon was adequate. This was quite a disadvantage in comparison with the first study, where every step of the testing phase had mirrored a step of the learning phase. Still, this is at the same time advantageous in respect of the research question: whether 'text reads text' when it comes to subtext.

12.5 The testing phase

12.5.1 Stanza 1

This section will describe the task related to the first part of the poem and the student's responses. In this initial phase of the research I endeavoured not only to test the students' understanding of subtext, but also to make them see the process of relexicalisation (as these two phenomena appear together twice in the poem). However, I practically abandoned relexicalisation as part of the research after all 3 students repeatedly failed to find relexicalisation in Stanza 1 (I will at a later stage mention an exception to this). The students had read about relexicalisation in Louw (2010b) and offered comments to the paper that made me think they understood the term. However, they kept finding wrong examples, although I offered several additional explanations and illustrations on email. Therefore, I decided to keep the second study focused only on subtext at this early stage.

The first task was to comment on Yeats's use of 'sought a' (line 1of the poem, see Section 7.2.1) after studying an attached concordance. The concordance comprised all 79

lines that had been found in the corpus of the *Times* 1995 edition. I assumed the students were aware it came from the *Times* and not a different corpus because this information was contained in the concordance, but I had omitted to check if this awareness existed. I also assumed the students would bear this in mind when giving interpretation.

In her response, Student 1 pointed out that 'sought a *' in the corpus is mainly related to the legal register (sought a court ruling/divorce/final legal opinion/ban/judgement/judicial review, etc.). To her, it looked as if the persona had made an official request and was now waiting for an answer. She also noticed that the first lines contain an invocation. In my response to Student 1 I urged her to bear in mind that the concordance had come from the *Times* corpus, and to give her interpretation of Yeats' syntactic usage of 'sought a'. She then commented on 'sought' acting as a participle in the corpus, forming perfect tenses, whereas Yeats uses it in the past simple. The student claimed she sensed it could influence the meaning.

Although Student 1 did not infer from the concordance what I infer from it in my 2014 chapter, she did see an aspect of the meaning of 'sought a' in the concordance that I do not mention: 'sought' in the given reference corpus can often be replaced with 'asked for'. The legal register may account for it, but the *Times* newspaper, or newspaper language generally, is not primarily devoted to court cases. Besides, she connected it, very wisely in my view, to invocation. From the point of view of the mechanics of the task (that was the concordance she had received), the response by Student 1 was original and insightful.

Student 2 drew my attention to the following lines:

> Winter and summer <u>till old age began</u>
> <u>My circus animals were all on show…</u>

Student 2 connected 'sought a' to the collocates found in the *Times* concordance, like 'new', 'remedy', 'solution' and 'stay'. She suggested that the persona's true source of despair might be old age, and not writer's block, and her main argument was that 'new' as a collocate from the concordance is the opposite of old and may be associated with what is young. This approach can be related to Louw's notion of *similar events* (which had never

been mentioned to the students), especially since the student referred to particular citations in the corpus.

It seems to me, however, that this interpretation gives too much prominence to old age. This does not undermine the interesting notion that sought objects are seen as a way out or as new solutions as opposed to old ones; rather, Student 2 may not have had, or given, enough reason to suppose that old age is the true source of concern for the persona, and not writer's block. She herself felt that her obstinate impression was less than well-founded and more of a hunch, as could be seen from her last sentence. Doubtless, there is a connection with old age in the stanza, but, in my view, she should have offered more justification than just the collocate 'new', appearing in the concordance 6 times out of 79 (7.6%).

Student 3 noticed a clash between the formality and even 'propriety' of 'sought' in the corpus (the contexts, in her view, described what was a proper course of action) and the nature of inspiration, which she thought ought to come spontaneously. This aspect of 'sought a' was new to me. Although it was my impression that Student 3 had not engaged with the reference concordance as deeply as the other 2 participants, she had a new discovery.

Of the 3 students, only one needed to be reminded of the source of the concordance lines being the *Times* newspaper, and even then she did not modify her response. The interpretation offered by Student 1 was, in my view, the most original and striking: she saw that 'sought 'often replaced 'asked for' in the concordance, and connected Yeats's writer's block with the age-long custom of invoking muses and asking them for inspiration. Although Yeats does not address anyone in the studied verse, this intertextual connection is highly valuable as well as original. The interpretation given by Student 3 was also original, in her remark that the persona may have approached writing a poem too formally. The analysis that came from Student 2 took into account the collocates of 'sought a' in the reference corpus, which seems the most detailed approach of the three. Her response was also the longest. Still, her conclusion was the farthest from the scientific and the provable: instead, she allowed herself to be prompted by a hunch which she then did not spend sufficient time on justifying. This combination of detailed analysis and emotional response

is most interesting from the point of view of developing a methodology for classroom (and student-centred) corpus stylistics approaches. One of the possible explanations might be that, being interested, Student 2 invested more effort – but, being biased in favour of a preconceived interpretation, she made her scientific findings (absent collocates) work to promote her bias. Interestingly, and contrary to my expectations, none of the students saw the implications that I discuss in Louw and Milojkovic (2014): the fact that the object of the verb 'sought' followed by 'a' is always specified, and it is highly unspecified in Yeats.

The other collocation I had found of interest in Stanza 1 was 'being but a broken man', which I had chosen to research through 'but a *ed' in my own analysis of the poem (see Section 7.2.1). All 3 students showed the basic skill of judging the implications of *all* the QPVs present in the concordance I offered. Student 1 and Student 3 saw the aspect of something that had been used or destroyed (and, therefore, a transition from the new to the old). Student 3 in particular emphasizes this, saying that all the 'positive connotations a notion once held are now gone'. These 2 students immediately proceeded to apply the QVPs to the poem. Student 2 (whose response to the first question was the longest but again the least justified by evidence) this time went further. She noted that all the QVPs might refer to clothes (showing a greater ability to generalize than the other 2 participants), and introduced intertextuality: she commented on another poem by Yeats ('Sailing to Byzantium'), where an old man is compared to a 'tattered coat upon a stick'. Ironically, Student 2 also claimed that 'charred' was the only QPV that *could not* refer to clothes – not seeing that the lexical collocation in the given *Times* context was 'a charred nylon cardigan'. It is as if one might, especially if lacking experience, overlook what is evident in an attempt to perceive a deeper meaning.

In sum, all 3 students used reference corpora to interpret the given sections of the stanza and showed some success. Two students introduced intertextuality to the task, without using this term and without being taught this possibility. One student in particular, while showing great interest, gave a slightly prejudiced response to the first question but the most insightful one to the second, referring not only to intertextuality in Yeats but previously showing how to generalize on all the 5 appearing QPVs. Paradoxically, this

same student missed an important and very obvious collocate in one of the reference corpus contexts.

12.5.2 Stanza 2

The second stanza begins with 'What can I but enumerate old themes'. The string 'what can I but' being rare, I offered the students 17 contexts found in the Google Books – Fiction corpus (they are represented as a concordance in Section 7.2.2):

1 My dearest Mammy, I must write you a line, and kiss my dearest old Mother, though we differ ever so much about the Old Testament. What a deal of heart-burning & unkindness what division between friends has that book caused! – It can't be otherwise with your views regarding it, what can you do but deplore the error of those who won't receive it – what can I, but **say** my say too, & trust in God if I'm wrong – Trust if I'm wrong? It would be mistrust & a sort of Atheism in me to doubt for a moment that He will be good to me and all creatures…
(Israel at Vanity Fair: Jews and Judaism in the Writings of W. M. Thackeray. Siegbert Salomon Prawer, Brill, 1992, p. 340)

2 Me he assails, a poor defenceless foe:
Thus prefacing the fight – (if 'tis a fight
Where I must tamely bear a ruffian's might: -)
He flops, cries 'stand:' what can I, but **submit**?
To fight a drunken bear, were want of wit.
(Juvenal, Satire III, from A Translation of Juvenal and Persius into English verse, transl by Edward Owen, 1786, p 25)

3 'Flee from the wrath to come,'
I hear Jehovah say;
What can I do – let doubt be dumb, -
What can I – but **obey**?
His sceptre or His rod,
Who shall control them? None:
All things are possible with God,
He speaks, and it is done.
(The poetical works of James Montgomery: in six volumes. Original hymns. 1860, p. 26)

4 I hear, most gracious Queen – and having heard –
So rich the spell of beauty and so strong
The lure of power; reason so full and just
Inspires my words – what can I but **obey**?
(Poems lyrical and dramatic, by W. G. Hole, 1902)

5 But I! – look on my new-born child! –
  Look there
On my young wife! What can I but **despair**?
(Specimens of Persian Poetry, from Fraser's magazine for town and country, 1839, Volume 20, p. 131)

6 Still in my breast the tyrant reigns,
  And still he triumphs there:
Thus, thus oppressed with endless pains,
  What can I but **despair**?'
(The London magazine; or, Gentleman's monthly intelligencer: Volume 7, p. 356, 1738)

7 She is as bright as daylight, none can with her compare.
I fade, I droop, as doth green leaf in summer's sunny air,
If all my thought me helpeth nought, what can I but **despair**?
(A history of English literature from earliest times to 1916, by Arthur Compton-Rickett, 1946, p/ 55)

8 And what can I now but bless thee with all the strength of my heart, and what can I but **pray** for thy whole blessedness, O gentle soul of sinless beauty?
(Hymn of the Divinity of the Soul, from In the Heart of the Holy Grail, by James Leith Macbeth Bain, Kessinger Publishing, 2003, p. 139)

9 But if Destiny to this consent,
What can I but **my soul in sorrow bow**,
With tearful eyes and face in sadness bent?
(After the death of Laura, Petrarch, tranls from Italian, from Original poems with translations from Scandinavian and other poets, by Sophia Miligan, 1856, p. 308)

10 Still, when your truthful eyes,
your keen, attentive stare,

endow the vacuous slut
with royalty, when you match
her soul to her shimmering hair,
what can she do but rise
to your imagined throne?
And what can I, but **see**
beyond the world that is,
when, faithful, you insist
I have the golden key -
and learn from you once more
the terror and the bliss,
the world as it might be?
(The Private Life: Poems by Lisel Mueller, LSU Press, 1976, p. 6)

11 It is not, What can I? but, ***What can't He***? as somebody says. Go on fearlessly. Trust Him through all. Work as hard as you like, only trust and be easy, and be sure then that the right way, and only right, will open to you, and that just at the right ...
The Quiver, 1880, p. 286

12 Naught can I but spill the wine of weeping and my garment stain.
How should 1 avail to draw one breath, nor like the flute complain?
What can I but, ***like the ended banquet, desolate remain***? '
'Fears are heedless, spheres are ruthless, Fortune is inconstant quite;
'Woes are many, friends not any, strong the foe, and weak my plight'

(A History of Ottoman poetry, Elias John Wilkinson Gibb, Edward Granville Brouwne, 1990, Volume 3)

13 Thy equal Nature never formed,
Dear maid, with every charm adorned;
With Hee’s youth and Venus’ grace,
And Cupid’s smile in thy sweet fece.
Thy beauties ev’ry heart subdue,
And conquering heros sigh for you;
And knowing this, ah! What can I
But **love** in silence, pine and die?
(The universal songster: Or, Museum of Mirth, Volume 3 (Google e-book)

14 Before he could get further, she continued, sinking lower at his feet: 'Ah me, my Lord, if now thou art thinking me bold and forward, and outcast from natural pride, what can I but **plead** the greater love I bear you as my benefactor and sovereign…'
(The Prince of India, by Lew Wallace)

15 *Doria* …the contact, love
With thy hair, polish'd stem, hath smoothed away
The knots and lichen from my bark of life,
And made it shine with brightness scarce its own.
What was I till I found thee!
Nina Do not cease!
Say on, say on!
Doria What can I but **repeat**
A tale already told?
(Nina Storza, a tragedy, by Richard Zouch R. Troughton, 1855, p.40)

16 So softly, purely glows Her face, earth has not smutched it
Since the essential rose
In some far silence touched it.
What can I but **repeat**
The vow of every mother –
There is not one so sweet
In all the world – none other!
(The Life and Death of Mrs Tidmuss: an epic of insignificance, by Wifrid Blair, 1923)

17 WISE MAN. Twice have I dreamed it in a morning dream,
Now nothing serves my pupils but to come
With a like thought. Reason is growing dim;
A moment more and Frenzy will beat his drum
And laugh aloud and scream;
And I must dance in the dream.
No, no, but it is like a hawk, a hawk of the air,
It has swooped down—and this swoop makes the third—
And *what can I, but* **tremble** like a bird?
FOOL. Give me a penny.
WISE MAN. That I should dream it twice, and after that, that they should pick it out!

(The Hourglass, from The Collected Works of W. B. Yeats. Volume II: The Plays, Simon and Schuster, 2001, p. 278)

Figure 12.1. All the contexts of 'what can I but' in the Google Books – Fiction corpus

In order to gain maximal advantage from the students' responses, the reader ought to be acquainted with the findings stated in Louw and Milojkovic (2014). A short recapitulation follows:

> All contexts can be grouped into 3 major themes: despair (5, 6, 7, 9, 12, 13, 17), faith in God (1, 3, 11, 17), love (5, 6, 7, 9, 13, 14, 15, 16), and submission (3, 4, 9 (to fate), 14, 17). New vision is strongly present in 10 and 17.
>
> The play 'The Hourglass', where Yeats himself uses the grammatical string in question, deserves particular attention [...] It is obvious that the usage of the grammatical string in 'The Hourglass' expresses despair, faith in God, new vision and submission. Love is not at all mentioned in the play, and therefore, cannot be considered as subtext of this line in 'The Hourglass' (Louw and Milojkovic 2014: 272-273).

In this case, the subjects also agreed in their interpretation, quoting resignation, pessimism and helplessness as semantic auras. Each student also stressed angles which were not mentioned by the other two. Student 1 could see that the action denoted by QPVs was not expected to make a difference, but was being done all the same. Student 2 counted the QPVs and noticed 'repeat', which is also Yeats' choice in context 17; her analysis seemed the most detailed and complete. To Student 3, the QPVs implied putting one's hopes into someone else, and she noted this was not what Yeats's persona does in the poem. I do not mention this aspect in my own analysis, and therefore this is the student's contribution to my understanding of the poem. My feeling is and was, however, that Yeats *would like* to put all his hopes into someone else, and, without Maud, the ladder, and the circus animals, God remains – which corresponds to the point I make in my interpretation that this grammatical string relexicalises the immediately preceding 'the Lord knows what'. It is worth noting that not one student commented on 'tremble' in context 17 being Yeats's

choice in ‘The Hourglass’ – they may not have focused at all on the sources of the corpus contexts.

It must be noted that, while generalizing on hopelessness, passivity and fruitless action, the students did not notice certain interesting particulars about the string’s subtext: that submission is likelier to be directed at God and that despair is likelier to be caused by love (not fate, as in Context 12). None of the subjects thought of connecting the subtext of this first line to the whole middle section or the poem as a whole, where thematically Stanza 2 is focused on love, Stanza 3 on faith and Stanza 4 on despair. Although the students did notice general trends which the QPVs of the given string point to, the fact remains that they could have come up with much more had they been aware of all the possibilities that subtext offers stylistic interpretation, and this can only come with experience.

12.5.3 Stanza 5

The poem consists of 5 stanzas and 3 sections; stanzas 2, 3 and 4 constitute the middle section. I hypothesise in my own analysis that ‘what can I but’ is subtextual not only for stanza 2, but also for the whole of the middle section.

The coda, of course, is the most effective of the three parts. In the coda, ‘now that my ladder’s gone’ seemed both crucial for interpretation and convenient for analysis. Therefore, as their last task, the students were asked to focus on ‘now that my * is’ and to study the data taken from reference corpora (see Section 7.2.3). Again, it seems useful at this point to quote my interpretation so that the reader has an opportunity to judge the students' analyses more objectively:

> These QPVs share one quality: extreme meaningfulness to the owner. If classified, there are two sub-groups:
>
> - a very significant person
> - an essential quality without which life or honourable existence is impossible (Louw and Milojkovic 2014: 279).

While correctly noticing the negative connotation, the elegiac tone, and the sorrow for something that has passed for ever, Student 1 focused, inexplicably as she herself admitted, on the variable 'father'. From this one variable, but also from the fact that is it the most frequent on the lists of COCA and the Google Books – Fiction corpus, she concluded that the persona must be viewing his poems as his children. This conclusion suggests a converse relationship, especially since Student 1 herself connected 'father' with support – how could the persona have felt the loss of support and be the loss as 'father' to the poems, given that 'father' is preceded by 'my' in the poem? Still, there are 2 variables on the list of the Google Books – Fiction corpus that fall into the category of children: 'son' (117) and 'daughter' (51), so this line of reasoning could have been pursued with more justification.

Unlike Student 1, Student 2 did notice the division between those QPVs that refer to family members and 'words like heart, mind, life', and noted that 'gone' can mean both 'left' and 'died'. Also, apart from noticing the aura of irrecoverable loss, Student 2 believed that the variable ‘father’ showed that the persona was no longer the father figure to the circus animals, similarly to Student 1. She offered more justification to support the claim, using the variable ‘master’, while preserving the opposite relationship: it is the persona who complains of the loss of the variable. She also connected the variable ‘father’ to ‘gone’ meaning ‘dead’, and to the variable ‘life’, and this gave additional support to her idea that in the poem the persona complains of old age and approaching death rather than of the loss of inspiration. She also noted the ‘dignified’ aspect of many variables in the context of the superior status of the ladder – not forgetting the strangeness of ‘ladder’ when viewed alongside the variables. She called this strangeness ‘opposition’, clearly because, to her, a ladder was too physical an object to be filling the slot usually occupied by what is dignified.

Student 3 gave a good critical account of the meaningfulness of 'ladder' as a 'means of achieving elevation', but, apart from the statement that ‘the word choice indicates that the poet is experiencing a loss of something important to him’, her interpretation was not corpus-based. When I prompted Student 3 to give more detailed feedback, she did dwell on inspiration being an abstract notion, and therefore corresponding to QPVs ‘life’,

‘innocence’, and ‘honour’, and also on the positive connotation and the general significance of what is referred to by the QPVs of ‘now that my * is’.

Overall, it seems that this grammatical string with its QPVs inspired the students to write more than usual. They all agreed on the main point, which Student 1 defined as ‘sorrow for something that is passed for ever’, Student 2 as ‘the idea that something is lost and someone (the poet) must reconcile with it once and for all’, and Student 3 as ‘a loss of something important [to the poet]’. Specifically, both Student 1 and Student 2 saw the importance of the word ‘father’ and connected it to Yeats being the father of the poems (Student 1) or the animals (Student 2, who noticed the presence of the QPV ‘master’). Neither saw the paradox of this – if Yeats says that his ladder is gone, how can we substitute the persona for ‘father’ and say that the persona *is* the father?

Students 1 and 2 became aware of the natural division of the QVPs into ‘family members’ and ‘abstract’ variables, unlike Student 1. Student 2 noticed the absence of ‘youth’ and the presence of ‘life’, and connected it to the idea that only she expressed, namely, that ‘gone’ might mean both ‘left’ and ‘died’: the persona, she claimed, might be afraid of death. While Student 1 and Student 3 connected the ‘ladder’ to means of achieving elevation, Student 2 noticed (not unreasonably) its lowly status in comparison with the string’s most frequent QPVs, interpreting it as self-irony on the poet’s part. In this, she most probably processed the relexicalisation in the stanza: she saw the ladder’s symbolic ‘superior status’, while it is at the same time a physical object in the lowly shop.

Naturally, I was curious to see whether the students would connect the QPVs in this grammatical string to the absence of Maud Gonne, as I do in Louw and Milojkovic (2014: 279). Although this did not occur, Student 3 came closest to this hypothesis when she wrote: ‘The concordances mostly contain words that represent one’s closest human beings, members of one’s family. That may indicate that inspiration is an essential part of the poet’s life, that it is as important as family. The other words usually represent abstract notions of life, time, innocence, honour and the like. Those words fit more into the same category as inspiration (abstract).’ In stating that inspiration is as important as family, the student almost arrived at the tentative conclusion that Maud may have replaced in Yeats’s life the family he never founded with her.

12.6 Relexicalisation as part of the testing phase

Relexicalisation (see Section 2.5.1 for a discussion of relexicalisation and 7.2.1 for relexicalisation in Yeats) was not properly grasped by any of the 3 students, except for one example. The basis of their knowledge of the term was Louw (2010b), and my explanation by email, followed by a pun as an illustration. Although all students knowledgeably commented on relexicalisation after reading Louw's paper, they failed to identify its example in Stanza 1 ('my circus animals were all on show'), quoting 'broken man' instead. Despite my repeated explanation, they seemed to think that they were to look for a word used delexically that could be used as a fully lexicalized one elsewhere. In this way, the fate of relexicalisation in the second study was similar to that of subtext in the first, and for very similar reasons: not enough time devoted to the phenomenon and my insistence that text must read text and that a minimum of explanation was to be provided. An explanation is of no value if it is not understood; checking understanding must involve practical examples. I believe the students' misunderstanding of the term does not disprove the principle that text reads text, but simply shows that the mechanism of relexicalisation takes longer to sink in.

On the other hand, Student 2 was the most persistent of the three, and finally came up with an illustration and explanation which may be accepted as correct: 'winter and summer till old age began'. 'Winter and summer', she claimed, was a delexical expression relexicalised by 'began'. I believe this explanation could be accepted.

Moreover, the student's account of searching for an additional explanation suggests that such a one ought to have been provided by myself. My explanation ought to have dwelt more on the connection between relexicalisation, collocation and the reference corpus, and on how exactly we claim scientific status for relexicalisation: that it is created by proximity alone. Also, the student mentioned that if she did not finally understand what was required of her, this might cause a loss of confidence. This comment shows what additional factors might influence such investigations.

12.7 Conclusion to the testing phase

As this section summarises the findings of the testing phase, it is worth repeating here the research question of the second study: whether text reads text for third-year students of English in subtext-related tasks, and whether the CPT-based methodology would prove successful. On the basis of these 3 students' responses the answer to the first part of the question is yes. They accepted the notion of subtext and continued to apply it to the examples they were given with as much success as could be expected, considering their lack of practical experience. As for the part of the research question related to the success of the methodology, it is clear that the subjects would have profited from more practical examples before embarking on the testing phase. However, it was vital at this point to preserve their right to accept or reject subtext as a way to read text.

My overall conclusion is that all 3 students were able to generalize on the basis of the given QPVs, additionally coming up with original connections that had escaped me while I was writing my section of Louw and Milojkovic (2014). Even though their analyses varied in the degree of detail or completeness, each student provided a reasonably detailed interpretation that deepens the reader's understanding of the poem. Student 2, in particular, was aware of the existence of *similar events*, even if her analysis may have been focusing on an insufficient number of QPVs (as may have been the case with 'sought a'), or did not take into account the linearity of context. Relexicalisation was abandoned at an early stage of the study until better opportunities arose for a fuller description and exemplification. Nevertheless, one student found an example that I accepted as valid.

The flaws in the subjects' analyses were, no doubt, the result of very little theoretical background and no practical experience. The greatest flaw is arguably the mistaken notion that the QPV 'father' in 'now that my * is' leads to Yeats's seeing himself as father to his poems or the circus animals (as stated by Student 1 and Student 2 respectively). This logical flaw (Yeats cannot have fathered himself) somehow undermines the very good impression which the subjects' other findings may have created. This also shows a great degree of overreliance on frequent variables on their part, when they should have – in the spirit of *similar events* – taken the context of situation into account. However, students have a right to mistakes, naïve ones included, particularly as, in the interests of

research, they have not received any training in subtext extraction whatsoever. Moreover, this error points to the necessity of interpreting subtext before dealing with lexis in the authorial text, as emphasized Louw and Milojkovic (2016: 190).

Some other errors were also the result of inexperience. In my view, Student 2 placed too much emphasis on particular QVPs on the one hand, and approached the poem with a preconceived impression on the other. Being aware of similar events in the reference corpus, she may have attached too much importance to some collocates of 'sought a' in the Times corpus, not to mention the mistaken attribution of 'father' in the last string. At the same time, she intended, at the outset, to prove that Yeats is concerned with old age more than with writer's block, which may have led her further away from existing and provable connections.

Not all analyses were as detailed as I might have expected. For example, Student 3 suggested that the last string's QPVs may be divided into family members and abstract notions only after she was additionally prompted. Generally, analysing QPVs was all that was necessary in the last task, but both Student 1 and Student 2 claimed that this string was the most challenging.

Nobody paid attention to the fact that the last context of 'what can I but' from the reference corpus came from a play by Yeats. However, I noticed that generally the subjects were aware of the importance of the sources of the reference corpus, so in this case they must have overlooked the source of Context 17.

Still, in research of this kind errors on the part of the subjects are only natural. Research demanded that they devote their personal time and effort to a field they were not well-versed in. It is their achievements, and not mistakes through inexperience, which should be taken into account in such research conditions.

12.8 Students' feedback

In the feedback phase, the students were to read the parts of Louw and Milojkovic (2014) pertaining to the sections they had interpreted, in order to compare their analyses to mine, and to comment on the differences.

In her feedback, Student 1 gave a competent analysis of the aspects of the grammar strings that she had noticed and recognized the aspects that she had not seen. She could also spot the parts where her analysis could be improved upon, both where it could have been more detailed ('what can I but') and where it could have been more general (now that my * is'). This shows that she would have greatly profited from more practical work.

The feedback that came from Student 2 was also objective and competent. She admitted to using the corpus as an 'auxiliary device' and relying mostly on traditional methods of literary analysis, as well as to deciding in advance that the true concern of the persona is getting old and not writer's block. It is natural to postulate a small hypothesis ('Is the poet subconsciously lamenting old age? Will the corpus prove it through frequent absent collocates and the context of situation?'), but Student 2, in retrospect, could see that her findings were not straightforwardly corpus-based. As for Student 3, in her feedback she mostly dwelt on the implications of 'sought a', and on what might be represented by 'the raving slut'. As for my other findings and conclusions, she expressed agreement.

Interestingly, all 3 students chose to comment on 'sought a' – it seems to be more provocative than I could predict, probably because there are very many angles to view it from, each angle affording a slightly different interpretation. This example alone shows that a corpus stylistic analysis is never complete and absolute, and that best results are achieved in a collaborative effort to study reference corpora – the only restriction being that any discovery needs to be verifiable in the corpus. The testing phase needed to be conducted individually, but I imagine that 'sought a' would have prompted a lively classroom discussion, or a basis for a small project. As for other grammatical strings, the students' comments showed that their analyses could have been more confident with a little more guidance and experience. Student 1 recognized that she should not have employed the QVP 'father' literally, while Student 2 understood with hindsight that her interpretation should have relied more on corpora and less on intuition. On the other hand, they both claimed that they would have profited from more input: 'Now that I have read the chapter and know more about Yeats's personal life, I prefer your interpretation' (Student 1), 'It would be a lot easier to do it now when I have seen how it is done' (Student 2). Besides, the students did

not have enough biographical background: even at the moment of sending the last piece of feedback Student 1 believed that Yeats's important figure was deceased.

The comment by Student 3 differed from the other two in that she was never self-critical, which is understandable because her analysis had never gone completely wrong and had never been too intuitive. As to her trying to guess what was meant by 'the raving slut', the feedback was not meant to be corpus-based, so her guesswork at this point cannot be disqualified.

All 3 students' feedback was encouraging in that, without any prompts from me, they could for the most part plainly see the shortcomings of their own analyses. They also stated that their analyses could have been better if they had been aware of the mechanics of interpretation employed in the 2014 chapter. Combined with their complete lack of experience and very little theoretical background, this should predict CPT even more success in the classroom when the other necessary conditions are met.

12.9 Conclusion

The aim of this second practical study was to continue the investigation of the first one described in Section 11, and this time to focus on the subjects' understanding and acceptance of subtext as part of Contextual Prosodic Theory. The two-fold research question posed in the first study was retained in the second: whether the 'text reads text' principle worked for Belgrade (this time 3rd year) students of English, and whether the CPT methodology suggested in the first study was to prove successful. The success of the methodology in this study, however, might have been slightly hampered by the first research question, in that the students had received only an absolute minimum of theoretical background and no practical learning sessions of any kind. They literally 'learnt through doing'.

The fact that there were only 3 subjects is a huge difference from the first study and does not greatly contribute to the study's validity. On the other hand, all the subjects worked hard and were personally motivated by a thirst for knowledge, unlike in the first study, where motivation greatly varied. Moreover, this provided opportunities for the researcher to process feedback in greater detail.

The testing was to consist of the subjects' analysis of parts of the poem by Yeats 'The Circus Animals Desertion'; the feedback entailed their reading my, not then published, analysis of the same sections in order to compare it with their own and to comment on the differences. Unlike in the first study, this time the testing involved no grading system of evaluation, and the feedback was not based on a questionnaire.

The testing yielded highly motivating results for CPT-based classroom stylistics. Not only did the students handle the concordances with some competence, but they also came up with original and valuable angles of interpretation. They showed the ability to generalize on the basis of QPVs and to connect them to the studied text. Sometimes they showed awareness of *similar events*, not always understanding that the patterns ought to remain linear when their contexts of situation are being compared. Generally speaking, this and other errors (over-reliance on (some) QPVs and substituting scientifically provable facts with a pre-formulated hypothesis) are as valuable to the classroom research of CPT as the subjects' successful linking of concordances to the poem.

The qualitative part showed that with more background and instruction the students would have performed better, which they themselves admitted. By the background I also mean the life of Yeats, as it is the macro-context in this particular case, and by instruction I mean either access to existing corpus stylistic analyses, or to guided classroom practice. Unlike in the first study, in the second the students were deprived of a practical learning session for the sake of preserving the purity of feedback in responding to the question of whether, to them, subtext read text. With this in mind, I can only state that they performed fairly competently.

One important result of the second study is the variety of offered interpretations. In the reference corpus, 'sought a' is used more in the sense of 'asked for' and is therefore reminiscent of invocation; it is too formal to be used in the context of a theme because inspiration comes by itself. 'Being but a broken man' has the subtext of clothing in a state of disrepair and leads the reader to 'a tattered coat upon a stick' from 'Sailing to Byzantium'. 'What can I but' is sometimes used by a persona who hopes to rely on powerful others, while Yeats is (ostensibly) alone. The analyses overlapped in the obvious

generalities but differed in many particulars, and rarely did the students come up with interpretations I offer in my chapter, even if prompted.

Since the testing phase, albeit not involving any actual grading, did attempt descriptive evaluation of responses, the students needed to restrict their communication with me and, until the feedback phase involving reading my analysis, did not get any feedback on their conjectures. Still, it is natural to imagine what a lively conversation would have ensued among them if such had been allowed. The subjects' responses may be viewed as pieces of an emerging puzzle. Had there been more participants, more insights might have surfaced. This reminds one of the quest for meaning through community discourse described by Teubert (2010: 239):

> The meaning of any unique discourse occurrence results from its intertextual links with all other discourse occurrences. But it is we, the members of the interpretive community, who have to decide what constitutes such a link. And it is we who have to select the links to follow up. While our starting point is the prejudices, the foreknowledge and the fore-judgements we have learned from authority, our dialogue with the evidence of the discourse will liberate us from these preconceived ideas and take us to new shores. So will the interaction between the members of the interpretive community.

It is clear that the subjects' interpretations contribute to one another and suggest a similar picture to that described by Teubert in the quoted passage – bar the actual interaction which was prohibited by research conditions. The three participants of the study have undoubtedly joined the discourse community discussing meaning in Yeats.

On the whole, one may conclude from the students' responses that they never doubted subtext as deeper meaning, to such an extent that they sometimes overrated the importance of particular QPVs and/or took them literally. They were also given at times to very free interpretations, either trying to protect their own hypotheses, or sharing their own hunches. As Student 2 put it, to her, the corpus was still very much an auxiliary device rather than the source of scientific findings. These two extremes – a greater reliance on

particular QPVs than is objective, and a greater scope of associations than is scientific – may represent future common classroom mistakes, and their prevention ought to be incorporated into a future curriculum. Also, attention should be paid to *similar events* and the context of situation, for the analysis not to get out of hand. The subjects had never been instructed in these notions. All in all, the students performed the tasks in good faith, and never insisted on an interpretation, but expected to be corrected. Any hunch may be proved or disproved by corpora, and it is not the students' fault that there was no scope in this study for the verification of all their hunches. In the long term, it is precisely in the opportunities for fruitful discussion, for the verification of intuition, and for learning to walk the borderline between science and non-science without stepping into forbidden territory, that we find the potential of corpus stylistics as a university subject.

13. Prospection in novel-writing

13.1 Introduction: research of prospection in short poems

'One of the primary assumptions of written language is that it is premeditated; the writer has had time to consider it, and is aware of what is to follow as he/she writes any particular sentence' (Sinclair 2004: 103). This bears upon Sinclair's view of prospection, which he considered to inhere in phrasings of sentences as their ability to lead the reader to 'expect something specific in the next sentence ' (ibid.: 88). He thought it a mandatory part of coherent written discourse, which indeed distinguished it from not so coherent spoken language. Although he mainly viewed prospection as interaction between the current sentence and the next, he conceded that several sentences might intervene before prospection is realised, and that such fulfillment may be 'tactically delayed' (ibid.: 90). As shown in Section 3.2.10.2, Toolan (2009) subscribes to these views and has offered proof that delay or absence of such fulfillment may be exploited by writers as both suspense and surprise, or its realisation postponed until the very end of a short story.

As explained in Section 3.2.10.2, Louw's understanding of the term 'prospection' involves corpus-derived subtext (the most frequent QPVs of a grammar string) of the initial line of a poem. Louw (2013) shows how the first grammar string ('the * is *ed with the * of') of 'God's Grandeur' by G. M. Hopkins ('The world is charged with the grandeur of God') contains QPV 'filled' in the second lexical slot and QPVs 'drone', 'smell' and 'sound' in the third. They create the impression of the activity of man polluting the perfection of the world, which can be related to the second quatrain of the last stanza (see Section 2.8).

Another such example is Larkin's 'When first we met, and touching showed', discussed in Section 4.5. The string 'when first we *' yields 'practice to deceive' (a quote taken from Walter Scott's *Marmion*), 'met' (twice) and 'reduced'. Larkin's skepticism is not in the QPVs as such, except the first, but in the reference corpus contexts of 'when first we', which are all about underlying motives in relationships. Yet another example is the poem by Brodsky *Он знал, что эта боль в плече* ('He knew that this pain in the shoulder') discussed in Section 15.2. This is the premeditated quality of written discourse that Sinclair was referring to, and the reason for using poetry as the source of examples in this

dissertation - a short poem is a thought-through piece without digressions, where the interaction of semantic auras is at its best as condensed meaning.

13.2 Prospection in students' essays

It was discovered, however, that other types of writing contained prospection in the initial line - namely, 350-word essays written by third year students of English, University of Belgrade in December 2013 (Milojkovic and Louw 2017). One such example, also described in Louw and Milojkovic (2016: 183-188) is worth quoting here in full.

The topic offered by the teaching staff was 'What constitutes academic achievement?' Katarina Zivkovic's essay focused on arguing that academic achievement consists of gaining knowledge as such, not for the sake of future job prospects or grades, and that gaining knowledge involves effort. It is therefore inconsistent on the part of some students to regard passing exams as a goal in itself, or certain lecturers as overly strict. Katarina Zivkovic's essay began with the following sentence:

> As our society switches from rural to urban, it is becoming a trend among Serbian youth to go to university.

In the first clause, the researcher chose to look at 'as our * *es from'. This string seemed crucial to the clause, and also practicable to research. However, the reference corpora available did not contain such a string, except without its last element – 'from'. Therefore, the string 'as our * *es' was searched in the BNC, GB – UK and GB – US.

The search of the BNC yielded 31 different strings, many of which contained adjective + noun plural combinations after 'our', such as 'younger colleagues', or 'own eyes'. These had to be excluded, as it is unlikely that they could be followed by 'from * to *'. According to Louw, parts of speech should not be taken into account, but they must be in order to see if the grammar string in question may be part of a longer one. These are the remaining BNC-yielded usages:

1. as our train rushes (1)
2. as our survey indicates (1)

3. as our society becomes (1)
4. as our score gives (1)
5. as our planet moves (1)
6. as our market dies (1)
7. as our knowledge increases (1)
8. as our knowledge changes (1)
9. as our hero avenges (1)
10. as our faith assures (1)

Of these usages, although they all consist of ‘as our’ followed by a noun and a verb with an ‘es’ ending, not all could be followed by ‘from * to *’ (e.g. a market cannot die from * to *.) The possibility of being followed by ‘from * to *’ will hinge upon the verb – the QPV in the second slot. Nevertheless, we may as well note that the QPVs are ‘train’, ‘survey’, ‘society’, ‘score’, ‘planet’, ‘market’, ‘knowledge’ (twice), ‘hero’, and ‘faith’; ‘rush’, ‘indicate’, ‘become’, ‘give’, ‘move’, ‘die’, ‘increase’, ‘change’, ‘avenge’, ‘assure’.

Still, before generalizing on these QPVs, it is useful to single out the strings that may conceivably be followed by ‘from * to *’. These are

1. as our train rushes
5. as our planet moves
7. as our knowledge increases
8. as our knowledge changes

‘Society’, the QPV used by the student, has disappeared. The QPVs to be discussed are ‘train’, ‘planet’, ‘knowledge’ (twice); ‘rush’, ‘move’, ‘increase’, ‘change’. Since the student is discussing the gaining of higher education as a consequence of the changes mentioned, the QPV ‘knowledge’ is of particular significance. The least adaptable context, the rushing of the train, may contribute to stress the speed of changes in society. More static strings from the BNC, such as ‘as our survey indicates’, did not stand the test of the string ‘from * to *’ added to them.

This is in itself a conclusive finding, pointing to the nativeness of the subtext of the student's string. However, since more data is available, it is worth assessing if more can be gained from the Google Books – UK corpus. This will hopefully illustrate that data from different corpora overlap – and so it should, given that grammatical strings change much more slowly than lexis and thus should not depend on the variety, register, or even time of use (within reasonable limits). This is the data taken from Google Books – UK (adjective+noun plural combinations have likewise been excluded, as well as the odd exception, like 'ideas themselves'):

1. as our knowledge goes (1079)
2. as our experience goes (889)
3. as our evidence goes (537
4. as our information goes (391)
5. as our knowledge increases (281)
6. as our author observes (168)
7. as our author states (157)
8. as our observation goes (125)
9. as our author does (123)
10. as our knowledge advances (105)
11. as our knowledge reaches (103)
12. as our observation reaches (98)
13. as our experience reaches (92)
14. as our author expresses (91)
15. as our author supposes (90)
16. as our Lord teaches (88)
17. as our Saviour teaches (74)
18. as our Lord declares (66)
19. as our knowledge becomes (61)
20. as our story approaches (58)
21. as our Lord expresses (56)

22. as our sun declines (55)
23. as our memory serves (48)
24. as our society becomes (46)
25. as our Church teaches (46)
26. as our planet guides (45)
27. as our understanding increases (45)
28. as our moon does (45)
29. as our knowledge enables (42)
30. as our information reaches (41)
31. as our author imagines (36)
32. as our skill advances (34)

Again, the possibility of a string to be followed by 'from * to *,' will largely hinge upon the QPV in the second slot. A scientifically reliable way of dealing with such a large concordance with varying QPVs might be to look at the first ten most frequent uses. This will give us an opportunity to generalize without brushing aside findings which we cannot reliably process. These are the first ten most frequent strings in the GB – UK corpus:

1. as our knowledge goes (1079)
2. as our experience goes (889)
3. as our evidence goes (537)
4. as our information goes (391)
5. as our knowledge increases (281)
6. as our author observes (168)
7. as our author states (157)
8. as our observation goes (125)
9. as our author does (123)
10. as our knowledge advances (105)

The list above was scientifically obtained; it seems representative of what follows after position 10, as all of the QPVs in both slots appear after this position, with the exception of

'evidence'. As we are exploring subtext, and not co-selection, and as these QPVs are most frequent, we may comment on the two lists of variables as constituting the subtext of the studied grammatical string. The list that follows takes frequency into account:

Table 1. List of the most frequent lexical variables in 'as our * *s' from the Google Books – UK corpus

1st lexical slot	**2nd lexical slot**
knowledge (1079; 281; 105)	go (1079; 889; 537; 391; 125)
author (168; 157; 123)	increase (281)
experience (889)	observe (168)
evidence (537)	state (157)
information (391)	do (123)
observation (125)	advance (105)

Although neither 'society' nor 'switch' is on the list of the most frequent QPVs, it is possible to generalize on the subtext of the string 'as our * *s' on the basis of the ten most frequent uses from the GB – UK corpus. The first lexical slot is occupied by words belonging to the same semantic field: 'knowledge', 'experience', 'evidence', 'information', 'observation', which may be defined as input for the human mind ('observation' is a way of gaining input rather than input itself). Alternatively, the most frequent QPV in the first lexical slot is 'author'. In the second lexical slot the QPVs create an impression of activity ('go' is the most frequent, also 'do') and augmentation ('increase', 'advance'). It is worth noting here that 'increase' and 'advance' are by necessity types of change. Alternatively, the most frequent QPVs are 'observe' and 'state' – which are indeed co-selected with 'author' in strings themselves.

When applied to the analysis of the student's sentence, the most frequent QVPs in the GB – UK corpus deepen the impact of the sentence. The use of the string is undoubtedly native, as is suggested by the QPVs 'society' and 'change' (appearing on the BNC list). However, it appears that, as the student discusses higher education and academic achievement, she uses a grammatical string whose QPVs are 'knowledge' (the most frequent in the first lexical slot in GB – UK) and 'increase' (the second most frequent QPV

in the second lexical slot, also supported by the presence of 'advance'). It is useful at this point to compare this finding with the results yielded by the BNC, also suggesting change (position 7) and increase (position 8) in knowledge (positions 7 and 8). As the point of the essay is that higher education is about knowledge and not gaining financial security, it becomes apparent that the subtext of the first clause deepens its interpretation – and prospects the content of the essay.

This example has been quoted here as the most clear-cut and convincing. Still, it must be pointed out that 31.5% of the subject of the study started their essay with a clause whose grammatical string contained prospection (Milojkovic and Louw 2017: 132).

13.3 Aims and significance of the study

Essays are relatively short. The most recent study of prospection in texts (Milojkovic in press) tackles novels as part of a bigger project to deal with genre at the level of collocation. This is a potential to CPT that has remained unexplored. An empirical study of this sort would uncover specificities of various genres that are invisible to the naked eye and might help to differentiate genres at the level of word combinations in Firth's context of situation. Since any text is a combination of its meanings, a genre would be researched not merely stylistically, but also at the level of content.

Given that the proposed methodology starts at the level of microcontexts and moves upwards to the general characteristics of a genre, the task would require teams of trained researchers and years of study. However, one may choose, intuitively, authors and texts that are consensually considered most representative, and look at their collocations, assuming that those might also be representative of an author/genre. The findings, of course, may only be of use up to a point. It must not be taken for granted that their sum automatically describes an author, let alone a genre. Still, the fact of the general consensus might be taken into account when *a priori* choosing representative texts. The research question, thus, could be posed as follows: how do the genres differentiated by the general public differ at the level of collocation, with all the consequences of its findings for the level of content? The empirical findings of this sort might, in turn, propose certain adjustments in the currently accepted genre-based classification of texts.

This section explores the differences between modernist and postmodernist writing, but does not go so far as to propose to define their characteristics. As a modest start, its aim is to illustrate how initial paragraphs of two different texts are constructed from the point of view of CPT, and in particular, how corpus-derived subtext and particularly prospection may be seen as instrumental in the impact they have on readers. The texts chosen as representative of these genres are Faulkner's *Absalom, Absalom!* and DeLillo's *White Noise*. In accordance with the empirical nature of CPT, the opening paragraphs of these novels will be studied 'from scratch', without previous assumptions of any sort.

These are the opening passages of Faulkner's *Absalom, Absalom!* and DeLillo's *White Noise*:

> From a little after two o'clock until almost sundown of the long still hot weary dead September afternoon they sat in what Miss Coldfield still called the office because her father had called it that – a dim hot airless room with the blinds all closed and fastened for forty-three summers because when she was a girl someone had believed that light and moving air carried heat and that dark was always cooler, and which (as the sun shone fuller on that side of the house) became latticed with yellow slashes full of dust motes which Quentin thought of as being flecks of the dead old dried paint itself blown inward from the scaling blinds as wind must have blown them *(120 words)*

> The station wagons arrived at noon, a long shining line that coursed through the west campus. In single file they eased around the orange I-beam sculpture and moved towards the dormitories. The roofs of the station wagons were loaded down with carefully secured suitcases full of light and heavy clothing; with boxes of blankets, boots and shoes, stationery and books, sheets, pillows, quilts; with rolled-up rugs and sleeping bags; with bicycles, skies, rucksacks, English and Western saddles, inflated rafts. As cars slowed to a crawl and stopped, students sprang out and raced to the rear doors to begin removing the objects inside; the stereo sets, radios, personal computers; small refrigerators and table ranges; the cartons of phonograph records and cassettes […] *(120 words)*

Being of the same length and at the same position in the novels, these texts qualify for this kind of analysis, despite the fact that DeLillo's paragraph had to be end-stopped in mid-sentence. The following sections contain a corpus-stylistic analysis of these texts, a comparison of the findings, and an attempt to connect these to some existing literary criticism.

13.4 Lexical features of the texts under investigation

It is usually the case that interpretation of lexical meaning within the framework of CPT involves investigating texts through the semantic prosodies of their lexis, as well as examining the features of reference corpus contexts produced by co-selection. Contrary to this rule, corpora will not be at this point resorted to due to the remarkable clarity of the stylistic function of the lexis in the texts under investigation. The two texts offered, of the same length and in the same position in both larger texts, display such obvious lexical features with such remarkable regularity and consistence that they allow for certain conclusions to be drawn without recourse to corpora. Let us look at adjectives, nouns and verbs separately.

The adjectives found in both texts are juxtaposed below:

long, still, hot, weary, dead, dim, hot, airless, cooler, yellow, dead, old, scaling (13)

long, shining, west, orange, carefully secured, light, full, heavy, rolled-up, English, Western, inflated, rear, small (14)

At this stage we must on no account separate the meanings of these words from their immediate context in the novel (i.e. forget that the first five describe a September afternoon, the next three describe a room, 'cooler' refers to the noun 'dark', 'yellow' describes slashes, the next two adjectives, 'dead' and 'old' refer to paint and 'scaling' refers to blinds). This is to say that we must not attach just any meaning to the words on the list. Still, it is obvious at first sight that 'dead' and 'hot' are repeated twice, as well as that the absolute majority of the adjectives, in terms of meaning, describe atmosphere in the physical sense and are also here to create a psychological atmosphere: 'still', 'hot', 'weary', 'dim', 'airless', 'cooler' contribute to both, while 'dead' (twice) and 'old' are significant

psychologically. We do not need a reference corpus to tell us that paint does not at any point of aging become ‘dead’, not to mention September afternoons.

The whole descriptive atmospheric and psychological effect becomes all the more obvious when compared to DeLillo’s list of adjectives. Though contributing to atmosphere in the psychological sense, in terms of meaning they describe physical objects, not surroundings or air. In particular, they tend to ascertain the exact nature, look and shape of objects described. As for objects themselves, the list of nouns will give us a clearer idea of what the adjectives in both texts refer to:

sundown, September, afternoon, office, father, room, blinds, summers, girl, light, air, heat, dark, sun, side, house, slashes, dust motes, Quentin, flecks, paint, blinds (22)

station wagons, noon, line, west campus, single file, I-beam sculpture, dormitories, roofs, station wagons, suitcases, clothing, boxes, blankets, boots, shoes, stationery, books, shields, pillows, quilts, rugs, sleeping bags, bicycles, skies, rucksacks, saddles, rafts, crawl, students, doors, objects, stereo sets, radios, personal computers, refrigerators, table ranges, **cartons**, phonograph records, cassettes (40)

Although there are similar numbers of adjectives in both texts, nouns are almost twice more frequent in DeLillo’s writing than in Faulkner’s. Evidently, they describe, in their majority, objects that are needed for a student’s existence inside the campus. There is nothing wrong with things, or packing and transferring them, and the fact that the writer chose to focus on them could simply mean that they describe a healthy twenty-year-old’s existence. A healthy mind resides in healthy body, and one would certainly prefer DeLillo’s abundance of sports equipment to Faulkner’s dusty habitation, reminiscent of Dickens’s decades old post-wedding shambles in *Great Expectations*. Common sense and intertextuality need not be arrived at by means of corpus findings. Still, the catalogue of things as opposed to a similar list of impressions in Faulkner indicates a difference in focus, DeLillo’s palpable worldliness challenging our impression of literariness of language and Faulkner conforming to it. Faulkner is concerned with impressions, and as such, they are bound to be individual (except air temperature). DeLillo catalogues things - objectively, impassively, precisely.

Interestingly, when it comes to verbs, more verb forms appear in Faulkner's excerpt than in DeLillo's. Nevertheless, juxtaposing the lists shows how Faulkner achieves the effect of stillness in his passage:

```
sat, called, called, closed, fastened, believed, moving, carried, shone,
latticed, thought of, being, dried, blown, must have blown (15)
```

```
arrived, coursed, eased, moved, loaded, slowed, stopped, sprang out, raced, to
begin, removing (11)
```

In DeLillo's text, all the verbs (except for 'loaded') are associated with movement. Not so in Faulkner's passage, where there are at least three categories of verbs:

a) cognitive actions: 'called' (twice), 'believed', 'thought of'
b) states of people and things ('sat', 'closed', 'fastened', 'latticed', 'being', 'dried', 'blown')
c) 'must have blown' is in between the two preceding categories because 'blown' here is an action but the epistemic modal suggests assumption rather than witnessing the action
d) actions performed by air ('moving', carried')

Therefore, although Faulkner's passage contains more verbs than DeLillo's, their meanings in combination create the effect of stillness, non-doing, given that the air is the only actual doer. They also support the notion that Faulkner's text has more to do with impressions, and DeLillo's has the effect of a report. Grammar also helps, as the only epistemic modal (referring to a past action) is contained in Faulkner's passage, while DeLillo routinely sticks to the past simple tense.

It is beyond reasonable doubt that the reference corpus would have confirmed the intuitive conclusions drawn in this section. For sequences like 'long still hot weary dead September afternoon' and 'dim hot airless room' no corpus analysis is necessary as the states of affairs are clear enough. The same applies to 'long shining line', 'sprang out and raced', and the list of objects intended for a particular use. Apparently, purely lexical influences in the studied passages are powerful enough to be viewed superficially. Lexico-grammatical collocations in both texts will be researched in the next section.

13.5 Corpus-derived subtext in Faulkner's excerpt

The first stretch of Faulkner's text that can reasonably be investigated for its grammatical subtext is 'they sat in what Miss Rosa still called the office'. It contains potentially productive search lines: 'they * in what' and 'still *ed the *'. The first consists of grammar words only; the second must begin with a lexical word, otherwise the search line '*ed *' will produce results that are too general and vague because of its high frequency in the language.

The search line 'they * in what' yielded the following QPVs in the COCA (I accepted verbs in the past tense, regular or irregular):

1. they **believed** in what (25)
2. they **were** in what (6)
3. they **lived** in what (3)
4. they **engaged** in what (2)
5. they **took in** what (2)
6. they **wrote** in what
7. they **surfaced** in what
8. they **succeeded** in what
9. they **stood** in what
10. they **stayed** in what
11. they **settled** in what
12. they **played** in what
13. they **bounced** in what
14. they **involved** in what
15. they **evidenced** in what
16. they **determined** in what
17. they **decided** in what
18. they **conversed** in what
19. they **came in** what

The most frequent QVP on the list is ‘believe’. Can it be assumed to carry the meaning of this grammatical string at a deeper level? Generally, there are two types of such connection that previous sections of this dissertation have shown: subtext either confirms the meaning we derive intuitively, or it adds a layer of meaning that can be confirmed by context clues. One subdivision of the latter is prospection: the ability of subtext to predict later developments in texts. The example of the string in question, in fact, fits the latter case.

When it comes to the QPV ‘believe’ (25 occurrences), the wider contexts in the reference corpus show that, from the narrator’s point of view, the belief in question may be correct or otherwise, but it is always fairly strong. As to the QPV ‘were’ (6), in four occurrences the purpose of the noun phrase starting with ‘what’ is hedging, e.g. ‘they were in what appeared to be the heart of a sun’. In the other two cases the sequence is part of a longer structure. Hedging is also the function of ‘in what’ in the contexts of ‘live’ (3) and ‘engage’ (2). The case of ‘took’ (2) should be disregarded altogether, since it is part of the phrasal verb ‘took in’. The QPVs occurring once, in their majority, appear in contexts of hedging.

It is significant that ‘believe’ appears 25 times, several times more frequently than the other QPVs. If applied to the novel in question, we must recall that it is contained in practically the first searchable grammar string of its whole text. Prospection, as explained above, predicts the development of thought further in the text (Louw 2013, Louw and Milojkovic 2016: 181-188). Apparently, this could be the first instance of such mechanism found within a novel. ‘Believe’ may be said to foreshadow the whole behavior pattern of Miss Rosa, her daily life, an unhealthy routine strictly ruled by her worldview, her ideas of the present and, more importantly, of the past. It is not what has taken place, but what she accepts as having taken place, that counts. Her habitation, as described, is a picture of her vision of what life has turned out to be. It has decayed because it was not allowed to progress naturally, while other people’s houses change as life changes. Furthermore, Miss Rosa’s whole life, as the novel will show, has been lived according to what she believes, not to what is or was. Intuition alone may also suggest that ‘they sat in what’ does allude to a system of beliefs rather than to a factual state of affairs, as hedging generally is a device

that is intended to be sensed by the recipient of information. On the whole, this could be taken as an instance of prospection, the subtext of the grammar string alluding to events that the reader will be informed of in the course of the novel. This is also a subtle means of characterization that is felt at the intuitive level.

The search line ‘still *ed the *’ underlying the sequence ‘still called the office’ is, strictly speaking, a lexico-grammatical collocation rather than grammar string proper. On some occasions it helps to narrow down the implications of a grammar string by including a significant lexical collocate, as will be shown below (it is particularly helpful in languages with developed inflection and flexible word order, such as Russian, see Милойкович in press 2019a). This is the frequency list for this search line:

1. still looked the same 8
2. still walked the earth 5
3. still covered the ground 3
4. still remembered the look 3
5. still roamed the earth 3
6. still roamed the streets 3
7. still supported the war 3
8. still walked the streets 3

The sequence ‘still looked the same’ (8) requires no explanation and perfectly fits its context – it is abundantly clear from the author’s description that Miss Rosa’s house has looked the same for several decades, without a thing missing or moved, apart from having grown older as a consequence of lack of intervention. ‘Still walked the earth’(5) is used in the reference corpus in the meaning of ‘still lived’ of a species of people who have disappeared but at the time referred to were still present. ‘Still covered the ground’ refers to *frost, water,and rubble.* ‘Still remembered the look’ means exactly what it seems and refers to a state of affairs in existence in the past. ‘Still roamed the earth’ (3) refers to dinosaurs in two contexts and men to one. ‘Still roamed the streets’ (3) contains collocates ‘looters’, ‘armed bands’ and ‘gangs’, and might, if one so wished, be connected to some of the war scenes in the novel, as well as the sequence ‘still supported the war’ (3) – this is not an

oversimplistic or overly optimistic summary of findings, but merely a statement of fact that a connection, even an implausible one, could be seen. 'Still walked the streets' (3) appears in two contexts instead of three due to repetition in the corpus, and describes desolated cities or districts where they used to be still populated at the time referred to by the narrator. Overall, the semantic aura of this line does not only point to Miss Rosa's not having changed, but also to the time of her life having taken place some considerable time ago.

Can one subconsciously encode a grammar string for it to foreshadow war scenes as well as other events that are less specific, more generalisable and therefore inviting a more plausible connection? We can only search the corpus and describe its data. Language is a mechanism still to be demystified. The fact remains that these, and not other, contexts were yielded by the corpus.

So far subtext has alluded to Miss Rosa's system of beliefs and to their not having changed, as well as herself and her house, whose unchangingness is practically the subject of the passage under investigation. The next search string, 'her father had *ed', again a lexico-grammatical combination, yielded the following QPVs:

died (66), insisted (10), worked (10), passed (8), wanted (8), asked (7), called (7), explained (7), lived (7), married (7), owned (7), purchased (7), warned (7), moved (6), planted (6), started (6), turned (6)stopped, talked, tried, used, walked (4), founded, forced, decided, betrayed, arrived , believed, allowed, invested, managed, molested, presented, raised, suffered, treated, vanished (3), accumulated, accepted, abused, committed, collected, convinced, created, dropped, designed, encouraged, explored, disappeared, fondled, hired, gathered, hoped, intended, joined, laughed, looked, mentioned, ordered, perished, picked, placed, planned, pointed, promised, provided, raped, remarried, reveled, seemed, served, settled, slipped, smiled, survived, teased, tied (2), hunted, homesteaded, hinted , helped, hated, handed, guarded, grimaced, glued, finished, fed, farmed, faded, destroyed, expected, exclaimed (1) (total 365).

The most frequent OPV, 'died' (66), 6.6 times more frequent than the next one, is hugely significant to the passage. In fact, the QPV 'passed' is followed by 'away' in four cases out of eight and therefore the total of the QPVs with the meaning of 'die' is 70, not 66. Miss Rosa's father is indeed deceased, his voluntary starvation in wartime influencing the future of his daughter who then went to live with her sister – upon which, in that house, she lived through the offence of her life, and this led her to adopt the life-in-death existence which we witness in the passage. The other QPVs in the corpus show 'father' as an active

agent, an authoritative figure whose actions in the past have influenced the present of the characters in the given contexts. Therefore, this could be yet another instance of a grammar string prospecting what follows, in our case, the account given in the novel of the key transitions in Miss Rosa's past.

The sequence 'with the blinds all closed' contains the grammar string 'with the *s all *ed'. In the four contexts yielded by COCA, four are negative (including a positive reference to tidiness in the context of recent death), one is positive (line 4), and one neutral because it is merely descriptive (line 5):

1" He reached inside and plucked one of them off its roost. When he brought it back down, **I was startled to see just how small it was, bodily**. With the wings all folded up, it was mouse-size. A baby mouse. " See that chipmunk stripe down its back? " Rey said. " **That's not a natural species**. It's a nu-bat. They're gene-gineered, like the house.

2 Her **pain** was so alluring to me. I stared at the pictures of her **depressed** bed with the sheets all bunched up and stained with her bodily fluids and dried up menstrual blood and the **psychic weight of psychic bedsores from not being able to lift oneself out of there**.

3 She knew she would find Alexis there, and **it broke her heart** as she walked into her mother's pretty pink satin dressing room, with the perfumes all lined up, and the hats neatly put on the shelf, and the shoes all perfectly arranged... **the shoes she would never wear again**. Edwina tried not to look at them, as **her own eyes filled with tears**.

4 " I've been learning, too, " Shadrach admitted sotto voce, a grin shining from his face. " Miz Eliza, she leaves books out for me, with the lessons all marked down. Sometimes she even leaves paper out for me and I write things for her and hide them under her pillow.

5 In an 1895 handbook on how to make lantern slides, A. R. Dresser informs his readers that " masks are sold by dealers... in various shapes, viz., Circular, cushion (a square with the corners all rounded); dome (a square, with the top corners only rounded, the bottom corners being left square); oval; a square or oblong.

6 That was when he noticed the magnification setting on the optical array: **In his panic**, he'd apparently **bumped it** to its highest level. Whatever he had seen must have been several astronomical units away. Of course, he couldn't tell how many now with the controls all discombobulated.

Clearly, the proportion of negative contexts suggests that the grammar string in question carries a negative semantic prosody. It is worth noting that in the two contexts of it in BNC and one in the Google Books US corpus are all negative. The impression is that of an

accident, inhibition or frustration (as with the collocates ‘folded’ and ‘bunched’ in the concordance). This is in accordance with Faulkner’s use, as well as with the whole life experience of Miss Rosa.

‘When she was’ is a grammar string known to have been studied by Sinclair (see Section 2.5.1). He found that, when it came to professional development, the contexts were always positive (e.g. ‘approached’, signifying excellent job opportunities), and that otherwise they were ‘hair-raising’ (e.g. ‘diagnosed’, ‘raped). This may be called the first instance of investigating corpus-derived subtext, prior to Louw (2010). ‘When she was a *’ yielded the following five most frequent nouns in the wildcarded slot (other parts of speech were excluded): 'child' (255), 'girl' (108), 'teenager' (86), 'kid' (74), 'baby' (59). As for ‘when she was a girl’, COCA contains 108 contexts of this expression. When the first ten were studied, four proved to be positive, four negative, and two difficult to classify. Admittedly, this is not a representative sample, but it is sufficient to show at least that human experience of this kind may be diverse.

The string ‘someone had *ed’ in the sequence ‘someone had believed’ called up 927 QPVs in COCA. These are the first ten: called (48), tried (44), turned (43), **died (40),** asked (38), pulled (34), dropped (25), placed (23), used (22), dumped (18). The most frequent one, ‘called’, is similar to ‘believed’ in the aspect of making a mental connection. In the remaining contexts the QPVs produced no surprises. What interests us here is the fourth most frequent one, ‘died’. After the grammar string contained in ‘her father had called’, death hovers once again in the passage, even if obscurely.

Predictably, the next searched string is ‘* was always *er’ in ‘dark was always cooler’. If we exclude pronouns ('there', 'it', 'she' etc), and leave only nouns in the first slot and adjectives in the second slot (and not e.g. expressions with ‘her’), what emerges is a positive semantic prosody, i.e. in agreement with the investigated textual segment. All contexts were opened. Some uses were found to have critical and negative implications, but overall this string also qualifies for a semantic prosody, this time positive, and is also in agreement with Faulkner’s positive use:

1. agreement was always better

2. bigger was always better (2) *(critical, not positive)*
3. blood was always thicker
4. Brandon was always quieter
5. Celia was always eager
6. Dick was always smarter
7. Dorothee was always happier
8. Eric was always talkier *(negative context, AIDS progressing and encouraging unbearable talkativeness)*
9. Esther was always eager
10. event was always larger
11. family was always together
12. father was always eager
13. Gary was always eager *(mixed but still positive context)*
14. Gipper was always dapper
15. goal was always narrower *(specific rather than black or white but still positive)*
16. grass was always greener (2)
17. Harriet was always prettier
18. Holman was always eager *(positive context)*
19. Jackson was always skinnier
20. JFK was always eager
21. laugh was always louder
22. Lily was always eager
23. line was always better
24. mother was always sicker
25. music was always broader
26. name was always cooler
27. Natalia was always faster
28. paper was always eager
29. passion was always steadier
30. Pickering was always eager

31. Randy was always faster
32. readership was always broader
33. Royal was always bolder
34. satisfaction was always higher (2)
35. seafood was always fresher
36. Smith was always eager
37. Sweetheart was always bigger
38. talent was always larger
39. team was always together
40. truth was always stranger
41. uniform was always cleaner

As for the sequence 'became latticed with yellow slashes', the string under investigation was 'became *ed with'. The list of QPVs follows, with those normally associated with difficulty or distress highlighted in bold. Variables appearing once have not been included.

obsessed (206), involved (153), acquainted (117), fascinated (103), associated (86), **infected (45)**, intrigued (35), **disillusioned (33)**, preoccupied (27), **frustrated (26)**, filled (26), identified (25), **disenchanted (23)**, **dissatisfied (23)**, **bored (23)**, concerned (20), consumed (18), infatuated (15), **saturated (15)**, affiliated (12), intertwined (12), overwhelmed (11), entwined (10), **clogged (9)**, covered (9), enchanted (8), **disgusted (8)**, reacquainted (8), **annoyed (7)**, **confused (7)**, **entangled (7)**, entranced (7), enthralled (6), **contaminated (6)**, imbued (6), integrated (6), **irritated (6)**, mixed (6), **infested (5)**, **littered (5)**, **tangled (5)**, connected (4), **clouded (4)**, **burdened (4)**, **afflicted (4)**, **charged (4)**, **crowded (4)**, **exasperated(4)**, intoxicated (4), occupied (4), **polluted (4**), equated (3), **disappointed (3)**, **disaffected (3)**, **drenched (3)**, infused (3), **inundated (3)**, invested (3), layered (3), merged (3), **outdated (3)**, **riddled (3**), absorbed (2), allied (2), coated (2), **cluttered (2)**, **conflated (2)**, **congested (2)**, **distracted (2)**, dusted (2), **enmeshed (2)**, **crammed (2)**, enraptured (2), fused (2), **fouled (2)**, flooded (2), frosted (2), impressed (2), institutionalized (2), interlocked (2), linked (2), **overcrowded (2)**, **overloaded (2)**, packed (2), populated (2), **seized (2)**, **shadowed (2)**, **soaked (2)**, **spotted (2)**, **tainted (2)**, tattooed (2), twisted (2).

Of course, a close corpus study might reveal other variables with negative meaning, or even modify the superficial intuitive impression created by the highlighted QPVs. Still, there is no denying that the most frequent variable, 'obsessed' (206), fits remarkably well into the

context of Miss Rosa's – in fact, obsession is the word. No matter how harshly she was treated, her life-long seclusion is a sign of obsession. Perhaps what she resented was the only offer of marriage in her life gone awry after three months of anticipation, rather than the actual humiliation of it being conditioned by the birth of a male heir. Be that as it may, the corpus points up the underlying semantic auras of the passage.

'Wind must have blown them' contains the grammar string '* must have * them' with two lexical variables. In COCA, one collocation appeared twice: 'God must have made them'. The other strings were too diverse for me to generalize. I resorted to the Google Books (American) corpus. At least the search would make more sense on a bigger sample. Strings containing lexis other than nouns in both wildcarded slots were excluded. On the frequency list below, all collocations come from the same source, with the exception of one:

1 The **boy must have read them**, as he and the Spirit crossed the threshold (Charles Dickens, 127 times).

2 It seems my **uncle must have left them** here when he went to America… (Autobiography of Benjamin Franklin, 127 times)

3 **Julia must have brought them** for you, Hedda (Henrik Ibsen, 63 times)

4 The retreating **troops must have roused them** (Washington Irving, 52 times)

5 'God must have made them' – the first one on the frequency list to appear from diverse sources, instead of one, 51 times

6 **Isaak must have uttered them** (Edwin Arlingtom Robinson, from the poem 'Isaak and Archibald', cited in various sources, 44 times)

It is clearly line 5 that must have the greatest impact on our interpretation. All other quotes come from the same source, albeit from different editions. They must have a bearing on interpretation simply because a frequently published author cannot help but influence the language in question. However, what we must rely on most is the collocation featuring the word 'God', as it comes from diverse sources and must therefore be considered the most frequent. 'God must have made them' is archetypal, primeval and intertextual. Here Faulkner's choice of 'wind' and 'blow' may be deemed significant. God makes; wind blows. God has a plan; wind is arbitrary. Miss Rosa's existence, founded on obsession and

resistance to change and ending in fire and annihilation, has been purposeless and destructive, as nature can be when left to itself.

Faulkner's subtext is undoubtedly powerful and if we restrict our comment only to the first QPV of every searched string, its results are nothing but telling. The subtext of the sequence 'they sat in what' proves to be 'believed', which is the most frequent QPV by far. The author's expression 'still called the office', if we retain the lexical word 'still', has as its subtext 'still looked the same'. Another studied phrase, also containing a lexical word, 'father', has 'died' as its subtext, the QVP again being by far the most frequent. 'When she was a' calls up states of childhood or adolescence. 'Became latticed with' refers to blinds but in the reference corpus the most frequent variable is 'obsessed'. Finally, the arbitrariness of Miss Rosa's fate rather than a reasonable divine plan is seen in the sequence 'wind must have blown them', whose subtext is convincingly revealed to be 'God must have made them'. That the influence of belief, father's death and obsession are crucial to Miss Rosa's life story is beyond doubt; that she has conserved her girly essence despite the passage of time is equally obvious. Faulkner's subtext and its powers of prospection may be said to be precise beyond belief.

The same can be said of the rest of semantic auras in his text. 'Still *ed the *' also calls up contexts featuring a species of beings that even then were rare and are now altogether history. The sequence 'with the blinds all closed' carries a SP of inhibition and frustration (auras in the corpus rather than lexical variables of grammar). 'Dark was always cooler', on the contrary, carries a positive SP, as in Faulkner's line. These semantic prosodies of grammar strings not only confirm what we can intuitively conclude about the atmosphere in Faulkner's text, but also prospect more of the same, as anyone who has read the novel can verify. A focus on the past and inhibition is what Miss Rosa's life is about, while the reference to the dark carrying a positive SP is reminiscent of the relief Larkin's persona feels with the advent of night (Milojkovic 2011a).

If we accept subtext as valid means of investigating the deeper meaning of a text and go on to summarise its findings, we may say that, besides its accuracy in prospection, it reveals states rather than actions, and belief, focus on death and obsession rather than

doing. This agrees with the anaysis of Faulkner's nouns, adjectives and verbs in the previous section.

13.6 Corpus-derived subtext in DeLillo's excerpt

As for DeLillo's passage, the same procedure will be followed: those strings that seem both crucial to the meaning of the text and practical in terms of corpus searches will be analysed for their subtext. The first string to be studied is '*s *ed at' in 'the station wagons arrived at noon', the opening phrase of the novel. This is of vital importance, as, according to Louw, prospection should ideally occur in the very first line. The table below shows the classification of the lexico-grammatical combinations yielded by the corpus. The cut-off point was 11 occurrences, and strings that contained different grammar words were excluded. There is a clear division into four types of context:

Table 2. Concordance of '*s *ed at' in COCA

	active measures	*higher education and scientific research*	*involving looking*	*involving armed participants*
1	*programs aimed* at 219	**researchers looked** at 93	*eyes stared* at 65	shots fired at 36
2	*policies aimed* at 112	**students enrolled** at 78	*eyes widened* at 35	guns pointed at 32
3	*efforts aimed* at 96	**papers presented** at 59	*girls looked* at 32	missiles aimed at 27
4	*measures aimed* at 84	**studies conducted** at 31	*eyes looked* at 26	attacks aimed at 16
5	*talks aimed* at 61	**samples collected** at 25	*boys looked* at 23	officers **arrived** at 28
6	*interventions aimed* at 59	**studies aimed** at 23	*Lucas looked* at 15	weapons aimed at 13
7	*initiatives aimed* at 56	**levels measured** at 22	*eyes glared* at 14	guns aimed at 12
8	*activities aimed* at 48	**courses offered** at 21	*James looked* at 13	officers stationed at 12
9	*strategies aimed* at 47	**studies looked** at 20	*eyes peered* at 12	soldiers stationed at 12
10	*laws aimed* at 44	**students scored** at 19	*fingers pointed* at 12	agents **arrived** at 11
11	*projects aimed* at 42	**lessons learned** at 18	*kids looked* at 11	soldiers fired at 11
12	*reforms aimed* at 41	**students looked** at 17		soldiers killed at 11
13	*bills aimed* at 30	**classes offered** at 16		officers looked at 11
14	*campaigns aimed* at 30	**tests conducted** at 16		
15	*products aimed* at 30	**students arrived** at 16		
16	*proposals aimed* at 29	**students placed** at 15		
17	*ads aimed* at 28	**programs targeted** at 15		
18	*programs offered* at 26	**colleagues looked** at 15		
19	*rules aimed* at 26	**studies presented** at 12		
20	*actions aimed* at 24	**variables measured** at 12		
21	*regulations aimed* at 23	**experiments aimed** at 11		
22	*services aimed* at 22	**experiments conducted** at 11		
23	*efforts directed* at 21			
24	*criticisms leveled* at 18			
25	*changes aimed* at 18			
26	*scientists looked* at 18			
27	*provisions aimed* at 17			

28	*process aimed* at 17			
29	*magazines aimed* at 16			
30	*articles aimed* at 15			
31	*negotiations aimed* at 15			
32	*operations aimed* at 14			
33	*approaches aimed* at 14			
34	*programs directed* at 14			
35	*discussions aimed* at 13			
36	*questions directed* at 12			
37	*sanctions aimed* at 12			
38	*protests directed* at 12			
39	*protesters gathered* at 11			
40	*steps aimed at 11*			
41	*books aimed* at 12			
42	*messages aimed* at 11			
43	*lawsuits aimed* at 11			
44	*guidelines aimed* at 11			

A group of combinations remained unclassified: 'legs crossed at' (41), 'interests existed at' (30), 'things happened at' (30), 'doors opened at' (25), 'patients treated at' (19), 'investigators looked at' (19), 'cars parked at' (17), 'games played at' (16), 'bus stopped at' (16), 'twins separated at' (15), 'kids laughed at' (13), 'arms extended at' (12), 'bus **arrived** at' (12), 'workers employed at' (12), 'standards specified at' (11), 'crowds gathered at' (11), 'guests gathered at' (11).

The first and most numerous category reveals the active SP of the string, as shown by the three most frequent variables: 'programs aimed at' (219), 'policies aimed at' (112), 'efforts aimed at' (219). This category received the label 'active measures' and accommodated every combination with 'aimed' in the second slot. It included contexts where action was not physical (variables 'arrived', as in DeLillo's text, or 'gathered' were few and not taken into account), but rather undertaken as implementation of plans to improve a situation. The second largest group of contexts narrows down action as measures to action taken in higher education and scientific research, the most frequent being: 'researchers looked at' (93), 'students enrolled at' (78) and 'papers presented at' (59). The fourth category, involving armed participants, contains a greater variety of lexico-grammatical combinations than the third, named 'involving looking' (13 to 11), but is less numerous. Its first strings appear 65 times ('eyes stared at'), 35 times ('eyes widened at') and 32 times ('girls looked at'), as compared to the most frequent combinations in the

fourth category, where the most frequent string come up 36 times (‘shots fired at’), 32 times (‘guns pointed at’) and 27 times (‘missiles aimed at’).

When it comes to content and not frequency, DeLillo's opening phrase ‘the station wagons arrived at’ describes physical action – still, action it is, so implementing measures as its subtext simply deepens, or broadens, the meaning of the string. It is interesting that physical action in general is rare in the first category of variables, and DeLillo’s chosen variable ‘arrived’ in particular is rare in the corpus. Still, it may be taken to prove my initial assumption that active doing, and not passive impressions, is what DeLillo’s work depicts most. The eye-catching and unique fact is that the second most frequent category involves higher education and scientific research, and DeLillo’s book is conspicuously about a college professor who also does his own research on Hitler. Armed violence generally is again all too easily connected with the violent scenes close to the ending of the book. As to looking, the novel opens with the main character, a college prof, observing students’ families arriving with their belongings.

These connections may seem eerie. Life is complex enough for one to be able to connect any two random things, if one so wished. On the other hand, some topics appearing in the book – for instance, fear of death, or obsession with the media or shopping – are not shown in the table. As the study of subtext is at its beginnings, it may be best to reserve judgement. Still, my experience tells me that it is a rare occurrence for a specific area such as higher education and scientific research – and these are so interrelated nowadays as to form a separate field of activity – to be so prominent in a concordance. To remain on the safe side, let us say that the active prosody of DeLillo’s first grammar string supports the lexis in his paragraph, investigated earlier, and prospects emphasis on activity in his book, as much as Faulkner’s grammar supports his lexis and prospects contemplation and atmosphere.

The next string to be researched is ‘*ed through the *’ in the phrase ‘coursed through the west campus’. The first ten most frequent QPVs in the corpus also create an aura of physical movement, which corresponds to DeLillo’s use:

1. **walked through the door 200**

2. **stepped through the door 77**
3. flipped through the pages 59
4. looked through the window 54
5. *dragged through the streets 51*
6. *slipped through the cracks 50*
7. *pushed through the crowd 46*
8. viewed through the lens 45
9. worked through the night 45
10. **pushed through the door 41**

It is interesting that the most frequent lines, at positions 1 and 2, and also the line at position 10, depict moving through a door, as if signalling that a transition is taking place. In the string at position 10 the movement is forceful, lines 5 and 7 appear in contexts of violence (wider contexts were opened), and line 6 contains an idiom and carries an aura of dissatisfaction with those individuals who have escaped notice of society generally or an organization in particular. Suffice it to say that the aura of physical movement in the corpus agrees with DeLillo's context of situation.

The next investigated segment is 'they eased around the orange I-beam sculpture' and its underlying grammar string 'they *ed around the *'. The cut-off point was 2, but even strings appearing once retain the aura of physical movement that we see in the first most frequent lines:

1. they looked around the room 5
2. they worked around the clock 4
3. they **travelled** around the country 3
4. they **walked** around the side 3
5. they searched around the car 3
6. they **milled** around the sideboard 2
7. they **moved** around the room 2
8. they **walked** around the pond2
9. they **disappeared around the corner** 2

10. they **walked** around the factory 2

The aura of physical movement is retained in the next studied line, '*ed towards the *', as part of the sequence 'moved towards the dormitories'. Nine contexts were taken into account, and 4 occurrences were the cut-off point. Interestingly, the door again appears in the second slot:

1. **turned** towards the **door 6**
2. **moved** towards the **door 5**
3. **turned** towards the wall 5
4. looked towards the **door 4**
5. looked towards the window 4
6. **turned** towards the window 4
7. **walked** towards the back 4
8. **walked** towards the **door 4**
9. **walked** towards the **house 4**

Moving towards the door in these contexts, however, always involves getting out of closed spaces, not approaching buildings. Thus, strictly speaking, only the line at position 9 corresponds to DeLillo's context of situation – approaching a building.

The next segment worthy of investigation is 'were loaded down with carefully secured suitcases'. It was studied as two separate grammar strings. Let us view each. The string 'were *ed down with' absolutely fits DeLillo's context (the cut-off point was 2 occurrences in the corpus):

1. were loaded down with 6
2. were weighed down with 6
3. were washed down with 3
4. were weighted down with 2
5. were tied down with 2
6. were pinned down with 2

As for the segment ‘carefully secured suitcases’, the search line ‘*ly *ed *s’ proved very frequent and called up diverse contexts. The richness of the implications that this grammar string might carry was such that is did not yield to meaningful analysis when it came to the textual segment under investigation. An attempt to search the same string with the preposition ‘with’ at the beginning did not change this picture.

The textual segment ‘cars slowed down to a crawl’, containing ‘*s *ed down to a’ yielded contexts which are clearly conditioned by the word ‘down’. It mainly either indicates direction, or points to diminishing quantity or a narrowing of focus:

1. steps led down to a 5
2. stairs led down to a 3
3. woods chipped down to a 1
4. years strapped down to a 1
5. allegations boiled down to a 1
6. analyses boiled down to a 1
7. barrels machined down to a 1
8. besiegers settled down to a 1
9. contributors tapered down to a 1
10. eros glided down to a 1
11. focus narrowed down to a 1
12. foundations thinned down to a 1
13. hairs tapered down to a 1
14. issues boiled down to a 1
15. noises quieted down to a 1
16. predictions dwindled down to a 1
17. shoulders flared down to a 1
18. sounds died down to a 1
19. teams whittled down to a 1
20. tiers flannelled down to a 1
21. treasures handed down to a 1

22. vulture glideed down to a 1

The sequence ‘to begin removing the objects’ contains the grammar string ‘to * *ing the *s’ and mainly supports the idea of activity, as used by DeLillo. The difference between the author’s use and the reference corpus is that in COCA the first slot is most frequently filled by the verb 'to be' (the cut-off point was 3 occurrences):

1. to be calling the shots 11
2. to be making the decisions 6
3. to be doing the things 5
4. to be putting the brakes 4
5. to be watching the monitors 4
6. to keep doing the things 4
7. to keep asking the questions 3
8. to **start** working the phones 3
9. to stop reading the newspapers 3
10. to stop using the terms 3
11. to be watching the news 3
12. to be working the phones 3
13. to be pulling the strings 3
14. to be enforcing the laws 3
15. to avoid offending the Saudis 3
16. to avoid stating the obvious 3
17. to avoid repeating the mistakes 3

The verb 'start' (synonymous with 'begin') appears once, 'stop' twice, and 'avoid' three times. However, the semantic aura of continuous (be *ing, keep *ing) activity is evidently prevalent in the concordance.

To conclude, in DeLillo’s text corpus-derived subtext prevalently shows action. In the first investigated grammar string it emerges as steps taken rather than physical action, although DeLillo’s chosen variable does exist in the reference corpus. The next three

grammar strings support the idea of physical action, as in DeLillo's text, with variables such as 'walked' and 'moved'. The next string's subtext corresponds to DeLillo's ('loaded'). The next grammar string yields a variety of states of affairs in the corpus; these are not easily generalisable and do not contribute to detecting a deeper meaning. Finally, the last two strings carry the subtext of reduction and of activity, respectively, and this corresponds to their respective contexts. It must also be stated that the very first string, apart from the theme of activity, prospects college life and armed conflict, while supporting the theme of looking that exists in the actual text.

13.7 The comparison of Faulkner's and DeLillo's opening passages

If we compare the subtext of Faulkner's opening passage to that of DeLillo's, it is the contrast that makes this comparison worthwhile. Faulkner's text is about mental states and a focus on the past, while DeLillo's description is about action, first taken generally, and then as physical movement, the latter in complete agreement with the author's context of situation. Bearing in mind that these writers' chosen nouns, adjectives and verbs (see Section 13.4) also suggest this, the results of this section may be considered reliable. These and similar findings may be used to draw conclusions about writing styles employed in these novels.

While in Faulkner's text corpus-derived subtext adds to textual interpretation and may even prospect events that the reader will be acquainted with in the course of the novel, in DeLillo's passage every grammar string, with the exception of the first, carries no additional subtext. The first string used by DeLillo, however, prospects the developments and ending of the novel with such remarkable clarity that it cannot be put down to coincidence.

On the basis of this comparison we may conclude that Faulkner, in his passage, is concerned with irrealis, with atmosphere, with describing images in one's head. DeLillo, on the other hand, is a reporter, describing events as they occur, mentioning concrete actions and concrete objects. In the final count, Faulkner is about impressions, and DeLillo is about creating a real world, palpable and precise, no matter how skewed it may turn out to be as the plot unfolds.

13.8 Implications for literary criticism

The difficulty of this analysis is that Contextual Prosodic Theory insists on being empirical. If we adhere to the 'bottom-up' approach, we are not to discuss the differences between modernism and postmodernism prior to the corpus analysis of the novels that are supposed to represent them, but must first analyse these novels from scratch. It is not within the scope of this section to proclaim what is or is not typical of both approaches. Still, there is no harm in connecting the empirical findings to existing literary criticism, or juxtaposing one with the other. Dismissing existing criticism would be as unwise as affording it supremacy over empirical data.

According to Kershner (1997), the rise of modernism may be connected to the development of science, although, according to him, art in this regard precedes physics. In the case of *Absalom, Absalom!*, these two developments were in fact almost simultaneous. Quantum theory and relativity theory were indeed made public around the turn of the century, but the research and breakthroughs continued to appear at least well into the 1920s (Kershner 1997: 58). Anyway, according to Kershner (ibid.),

> […] it is tempting to find analogies between post-Einsteinian physics and novelists' experimentation with radically compressed and rearranged chronology. In fact, space and time are already interconnected for the novelist. The careful structuring and patterning of modernist novels makes them less an art form to be experienced entirely chronologically, as a sequence of events, and more an example of […] 'spatial form' – a work of art that must be visualized simultaneously in its entirety, as if it were a painting.

Interestingly, there happens to exist a connection between Contextual Prosodic Theory and quantum theory. Louw (2017c, see Section 2.9) actually compares subtextual findings to those related to quantum: these are forces that are not visible to the naked eye and no one had any inkling of their existence until they came up in measurements. Louw regrets that the developments in linguistics are behind those in physics and states that '[a] quantum subtextual inductive search of individuals' language carries with it […]

probabilistically determined "next moves" [which] become potential prospections, and […] the *revealing of hidden detail as objects or even personal obsessions*' (Louw 2017c: 90, my emphasis).

Hidden detail as objects and personal obsessions are indeed revealed in Faulkner's subtext. Prospection is clearly present also, and, as can be inferred from findings discussed in the previous section, to a considerably larger extent than in DeLillo's passage. Can this be explained by the modernist presenting of narratives as co-existent and as existing simultaneously with the present? If, in Faulkner's mind, the past and the present are rolled into one and co-inhabit life's landscape, it is no wonder that his heroine's past is inherent in her present as she steps onto page one of his novel. Prospection, thus, has been shown to act at the level of micro-contexts because it is realised as the future of the novel's text and simultaneously as the past of the novel's heroine. Time turns into space in the sense that it partly stops being chronological.

'Among the unresolved contradictions of representation in postmodern fiction is that of the relation between the past and the present', writes Hucheon (2001: 70). This, of course, describes texts which intertwine historical fact and fiction, and cannot be said of DeLillo's novel, as he creates his fictional world through the means of strictly factual, chronological narration. The story is told straightforwardly as fact. There are no clashing perspectives of characters on what goes on, or if there are, they are not different from real life, where people develop different views on current events. There is nothing of Faulkner's subtle juxtaposition or complementation of points of view. In this sense, Duvall (2008: 7) rightly calls DeLillo's America in *White Noise* 'dehistoricised'. This crucial distinction, noticeable in his writing overall at the intuitive level, is also visible at the level of corpus-derived subtext. Activity (as opposed to passive contemplation) is revealed in the subtext of the opening line of his text. After that, apart from prospecting active measures, college life and armed violence, the text of the opening passage becomes transparent: physical motion and transition is described in the actual text, and physical motion and transition it is when it comes to subtext. Reality has no alternative. Which means that, should reality of *White Noise* dissatisfy you – and it does – reality offers no escape.

DeLillo's *White Noise* 'is generally considered a prime example of postmodern fiction', according to the editors of Maus (2001: 45). They quote the opening excerpt of the novel as demonstrating Umberto Eco's claims about the importance of irony in postmodernist literature. Eco notes that 'with the modern, anyone who does not understand the game can only reject it, but with the postmodern, it is possible not to understand the game and yet take it seriously. Which is, after all, the quality (the risk) of irony' (2001: 46). After the opening line, there is literally no depth in DeLillo's passage, no corpus-derived subtext, and thus, perhaps, the irony is complete. Clearly, at the level of collocation the world he builds is a mere representation of the ordinary, as if he constructs his world by piecing it together from selected segments of the real one, rather than constructing a different world. This world may mean to shock – for example, college life in the opening passage of *White Noise* is presented as if composed merely of things – but as collocation, it is a perfect, unambiguous world. His means of narration is juxtaposition through the absence of one of the entities contrasted. Look at the world around you and look at DeLillo's scarecrow. If Jack the narrator has been referred to as 'a modernist in the postmodernist clothing' (Olster 2008: 89), DeLillo, stylistically speaking, is anything but a new Faulkner. He is said to have stated: 'I don't think of language in a theoretical way' (Cowart 2008: 161). Cowart goes on to say that 'DeLillo seems never to have been particularly daunted by twentieth-century ideas about language that… emphasize the problematic link between the signifier and the signified' (ibid.: 161-162). In the quoted passage at least, at the level of collocation, the two often merge in DeLillo's writing. The absence of prospection means that there is no potential for future developments in the text, and this means an absence of future. The absence of corpus-derived subtext that is distinguishable from actual text means that there is no alternative reality, not even as a possible world.

In this, DeLillo is so convincing that his irony has duped even literary critics. Referring to the theme of Hitler studies in the novel, Cantor quotes Bawer accusing DeLillo of insufficient dislike of Nazism: 'DeLillo's offense, to my mind, is that he refuses to make distinctions. To him, as to Jack Gladney, the question of Hitler is simply "not a question of good and evil" [...] Before rushing like Bawer to condemn DeLillo for moral indifference to

Nazism, we ought to recognise that *White Noise* in itself is not an example of Hitler studies, but rather a novel which portrays a professor involved in Hitler studies' (1991: 40). Cantor calls this narrative technique ‘routinisation’ (ibid.: 41). As for the passage that opens *White Noise*, Cantor says: 'With tongue in cheek, DeLillo suggests how attenuated America has become a community, how little holds it together as a nation […] All they have to unite them is a common culture, reflected in this communal rite of passage' (ibid.: 49). Still, to enforce this point, Cantor quotes the end of DeLillo's second paragraph, and not the beginning of the first, discussed in Section 13.6. And the second paragraph states the point plainly enough: 'This assembly of station wagons […] tells the parents they are a collection of the like-minded and the spiritually akin, a people, a nation'.

At least at the beginning of *White Noise*, the text states what it does. Lettrucchia (1991: 89) calls this narrative style 'phlegmatic'. Later he adds: 'he enjoys passing judgement dryly [...] the narrator is tired […] Nevertheless his weariness is disturbing because it is rooted in the banality of everyday life […] Gladney’s ironies bespeak boredom, an enervation of spirit, and, beneath it all, an intimation of a Hell whose origin lies in the high modernist fear that the life of consciousness, especially its aesthetic possibilities, will become like the mass culture that modernists may loath, often parody, and sometimes even love' (ibid.: 94-95). Innocence rules in paragraph one, and that is official. If some critics have condemned the novel as a 'schematic view of modern America', or 'lifeless', 'constructed upon a simpleminded political cliché', presenting life in America as 'dehumanized' (Ruppersburg and Engles 2000: 7), their views must have been affected by, among other things, the features of DeLillo's writing discussed in Section 13.6. However, even DeLillo, dry, phlegmatic and postmodernist in his absolute imitation of reality, produces a first grammar string fraught with implications of not only activity generally, college life in particular and observing as the narrator’s role, but also of the presence of an armed conflict, thus prospecting the ending of his novel. As Tony Morrison puts it, 'I always know the ending of my novels because that’s part of the idea, part of the theme'. Or, better still: 'I always know the ending; that’s where I start' (Morrison 1994: 101).

13.9. Conclusion

Although the corpus stylistic analysis of the opening passages of Faulkner's *Absalom, Absalom!* and of DeLillo's *White Noise* was primarily meant as an attempt at initiating research into modernist vs. postmodernist writing from the point of view of collocation, it seems to have done a greater favour to corpus stylistics and Louw's Contextual Prosodic Theory than to this initial aim. The findings of the study show not only that corpus-derived subtext, a new development in CPT, is an objective phenomenon that assists in understanding more fully the meaning of any lexico-grammatical collocation, but also that it contains prospection, which is the ability of such collocation to foretell later transitions in texts.

More precisely, in both paragraphs the first employed grammar string prospects quite clearly the themes that will be subsequently developed in the story line. After the first grammar string, however, the features of such collocations in the texts under investigation diverge. While in Faulkner's passage the text continues to prospect the events in Miss Rosa's life of which the reader is going to be informed at different junctures throughout the novel, DeLillo's subtext becomes strictly factual, corresponding to the text in the slightest detail in the majority of collocations. This is the first instance of subtextual prospection being found in the first lines of a novel.

If these texts may be considered representative of both novels, Faulkner's style emerges as much richer in implications, fraught with subtext as much as it consists of actual text. The lexis he uses, as well as the grammar, highlights an impressionistic, contemplative approach to reality, where text assumes the role of a signifier prospecting the signified. Irrealis is Faulkner's forte. DeLillo, on the contrary, is a reporter to the last. After the initial grammar string, the text does not prospect, and does not even for the most part contain subtext distinguishable from the implications of the actual text.

This section shows that these features have been highlighted by literary critics. In modernist writing, time has been compared to space as it co-exists rather than passes. Not only is it in complete accordance with Faulkner's ability to encode his text with prospection, but it is also in accordance with the character of Miss Rosa, whose life has stood still for several decades – which is precisely the major implication of the passage. As

for postmodernist novels, DeLillo cannot be wholly representative of this movement on the level of collocation. The deployment of CPT so far has shown that it is a rare case with a literary text not to contain deeper subtextual meaning – the previous studies connect this with inspiration (Louw and Milojkovic 2016, see also Sections 9 and 10). In Section 9.3 Lodge was accused of being banal when depicting a stressful situation using language that does not hint at emotional fluctuations. This is obviously not the case here – rather, the corpus stylistic analysis of DeLillo's writing shows that he does not play with realities or truths, but depicts only one existing unwavering reality. This may have contributed to his writing being seen as "dry", "phlegmatic", and even "lifeless".

While the analysis may have highlighted some of the stylistic features of the writers in question, it has not fulfilled the initial promise to delineate different writing styles of the two literary movements. Still, it shows the potential of corpus stylistics in reconstructing meaning in texts.

14 Corpus-derived subtext and the context of culture

14.1 Introduction

As was stated in Section 2.2.3.2, CPT studies collocation as part of Firthian context of situation. Context of situation, however, must always be considered within its cultural context, as was emphasised in Section 2.6.2.1. Louw in his thinking has always placed Malinowskian context of culture above Firthian context of situation (see Louw and Milojkovic 2016: 44, 53, 155-165). The reference corpus in CPT is viewed not only as a representation of the language, but also as a sample of the world (Section 2.9), thus affording the researcher an empirical opportunity of passing judgement on states of affairs in the contexts under consideration and classifying language events in the corpus.

CPT relies on the connection Malinowski made between language and culture because of Malinowski's collaboration with Firth (see Louw and Milojkovic 2016: 156-159):

> Malinowski's context of situation is a bit of the *social process* which can be considered apart and in which a speech event is central and makes all the difference [...] The context of situation is a convenient abstraction at the *social level* of analysis and forms the basis of the hierarchy of techniques for *the statement of meanings* [...] There is an element of habit, custom, tradition, the element of the past, and the element of innovation, of the moment, in which future is being born. When you speak you fuse these elements in verbal creation, the outcome of your language and your personality. What you say may be said to have style, and in this connection a vast field of research in stylistics awaits investigation in literature and in speech (Firth 1957: 182-184, my italics).

Although the context of culture was not explicitly mentioned in the previous sections, it was always understood as part of the analysis. The aim of this section is to show that the role of the context of culture in Contextual Prosodic Theory is twofold. On the one hand, it assists in textual interpretation by alerting the researcher to the relevant aspects of the context of situation. On the other hand, corpus stylistic research often leads to discovering specific detail regarding the context of culture in which the text under discussion

originated, be it a poem by Yeats or an excerpt from a novel by David Lodge. This two-way process of clarification does not easily yield to generalisation, and this is why it will be argued in this section that theoretical premises, such as Gricean maxims, are not required for a corpus stylistic analysis to yield valuable results, just as it is stated in Firth's passage quoted above (and contrary to the opinion stated in McIntyre 2018, see Section 3.2.9).

14.2 Context of culture as a tool in a corpus stylistic analysis

Louw (Louw and Milojkovic 2016: 56-59) illustrates the link between collocation and Malinowskian context of situation by co-selecting 'position' and 'untenable' in the Bank of English. His analysis confirms Sinclair's ideas of how language models experience (Sinclair 2004: 53), how 'an inherent likelihood of an event has to be related to the frequency of its occurrence in order to determine its linguistic role' (ibid.: 29), and how 'words cannot remain perpetually independent in their patterning unless they are very rare or specially protected [...] Otherwise, they begin to retain traces of repeated events in their usage, and expectations of events such as collocations arise'. These are reasons why Sinclair thought that a word on its own should be replaced by a 'largely phrasal' unit of meaning as the focus of linguistic inquiry (ibid.: 30).

After studying a concordance of 'untenable' from the Bank of English corpus (see also Section 2.6.2.1), Louw infers the following 'reliable rule':

> In the case of *men* whose position is declared to be untenable, most men refuse to resign and often survive the crisis. In the case of *women,* they either resign immediately or refuse to resign and are then *removed* by men (Louw and Milojkovic 2016: 58).

A relatively recent case of a position made untenable described by Milojkovic (ibid.: 159-161) can serve as an illustration of the role played by Malinowski’s context of culture. In November 2014 Emily Thornberry, a senior Labour MP and the shadow attorney general, representing Islington South and Finsbury, was campaigning during a by-election in Rochester and Stroud. During the campaign she tweeted a picture showing a semi-detached house draped in St George flags, with a white van parked outside. The text under

the photo read: 'image from #Rochester'. Within hours, Thornberry apologized to the voters in case she had caused offence and stepped down. A barrage of criticism ensued, to the effect that the tweet was 'derogatory and dismissive of the people'.[46] The tweet was called 'patronising and jaw-droppingly condescending' by the then Liberal Democrats leader Nick Clegg.[47] The then Labour leader Ed Miliband, who reportedly had personally influenced Thornberry's resignation, stated that 'her tweet conveyed a sense of disrespect'. The then Prime Minister David Cameron himself did not miss an opportunity to comment: 'Effectively what this means is that Ed Miliband's Labour party sneers at people who work hard, who are patriotic and who love their country, and I think that is absolutely appalling.' [48] However, in the surge of reactions there surfaced voices claiming that it was difficult to explain to non-residents of Britain what was so offensive about the tweet: 'Tried to explain to some Americans… Lost in translation. Stars and Stripes on many homes.' '…In Spain politicians often refuse to resign even when they are under criminal investigation'. [49]

It will be useful at this point to delineate the distinction between the context of situation and the context of culture:

> 'Context of situation' is what an observer needs to know about the immediate situation in order to understand a particular instance of language; 'context of culture' is what an observer needs to know about the broader culture in order to understand the meaning of what is being said or written' (Lee and Poynton 2000: 4).

The context of situation in the given example was easy to understand: at a crucial time the Labour party thought it wise to denounce an MP who appeared to 'sneer' at working classes by posting a given tweet. Certain elements of the context of culture were also understood by many non-residents of Britain: the working class traditionally voted Labour; the newly-emerged political party Ukip, which had succeeded in attracting some of working class voters (as well as some Tory MPs) had won the Rochester by-election; in

[46] http://www.theguardian.com/politics/2014/nov/20/emily-thornberry-resigns-rochester-tweet-labour-shadow-cabinet
[47] http://www.theguardian.com/politics/2014/nov/24/emily-thornberry-tweet-clegg-ukip-farage
[48] theguardian.com/politics/2014/nov/21/labour-founded-for-working-people-miliband-tells-white-van-man
[49] http://www.theguardian.com/politics/2014/nov/21/emily-thornberry-resignation-explain-outside-britain

these circumstances Ed Miliband needed a strong statement of support and respect for working classes represented by the owner of the white van in the photo. Still, even knowing all this, many remained baffled as to how an MP was sacked over a photo, including a Guardian author who claimed to have lived in England for 25 years.[50]

The mystery is resolved if one has absorbed the particulars of Britain's stratified society. Thornberry was MP for the posh Islington; St George flags are associated with working classes; these flags are sometimes associated with Britons of nationalist convictions; Ukip as a political party had expressed very many views supported by nationalist members of working classes who disliked the presence of migrants in the UK. By simply singling out a house decked out in St George flags (the working-class white van contributed as a visual collocate, as well as a rubbish bin and some litter under the window to the left), the Labour MP for posh Islington was seen as conveying contempt.

How can we, then, connect this event in the context of culture to the language event inherent in it? The caption under the photo said: image from #Rochester. Many commentators drew attention to the fact that the caption is neutral and not at all indicative of any reaction or message supposedly conveyed by Emily Thornberry. What follows proves and illustrates that the meaning of every instance of language use is determined, among other factors, by the semantic aura of the grammatical string inherent in it.

As CPT has not yet started dealing with the language of tweets and captions, it might be useful to study the grammatical string that would surround the message in the majority of registers: an image from Rochester. The string 'an * from *' was searched in the Google Corpus – UK. In the first 500 lines yielded by the corpus, all lines containing different grammatical strings (e.g. an article or an adjective following 'from') were excluded. The most frequent combination turned out to be 'an angel from heaven', sometimes with capital letters. Since language events in CPT are viewed in the framework of Firthian context of situation, the next step involved selecting only those lines in which 'from' was followed by a noun denoting a specified location ('heaven' is hardly a specified location). The data remaining from the initial concordance of 500 lines are presented below:

[50] http://www.theguardian.com/commentisfree/2014/nov/26/britain-stop-obsession-social-class

1. an **invasion** from France (594)
2. an **absence** from home (305)
3. an **invasion** from abroad (251)
4. an **exile** from France (227)
5. an **absence** from England (204)
6. an **oracle** from Delphi (201)
7. an **hour** from London (165)
8. an i**nvasion** from Scotland (163)
9. an **invasion** from Spain (162)
10. an **attack** from France (161)

Lexemes occupying the first lexical slot are highlighted in bold. The most frequent is ‘invasion’, followed by ‘absence’, followed by ‘exile’. Had Emily Thornberry used this exact grammatical string (as things stand, with the indefinite article missing, her exact text is practically stripped of an aura, in accordance with the register of captions), we would be justified in noting a certain negative logical SP ‑ corpus-derived subtext.

The analysis of the context of situation and culture presented above does not require any 'discussion of social or political theory' (McEnery and Hardie 2012: 141). It was not necessary to import Gricean maxims, nor any other instrument of a pragmatic or other theory in existence. Generally speaking, this type of analysis, reliant on the knowledge of cultural norms of a certain society, must not be dismissed on the basis of the researcher's 'non-corpus-informed, intuitive' perception of events (ibid.). Any such analysis must be thoroughly dissected as to its assumptions regarding cultural values of a certain society, and only then be accepted or rejected by a community of scholars (Teubert 2010: 239). The contextual element of a corpus stylistic analysis is 'non-corpus-informed' by definition, but an agreement as to its objective value can nevertheless be reached. Corpus data may confirm its implications, as in the example above.

14.3 A corpus-attested view a context of culture

As stated in Section 14.1, CPT may be employed to assist in the analysis of the context of culture in which a particular text emerged. This section, based on Milojkovic (2017b) will illustrate how such an analysis may be possible in practice. If cultural contexts are viewed as sets of stereotypes reflected in language use, and the reference corpus of a language may be taken as consisting of frequent and less frequent lexico-grammatical combinations, then a particular set of cultural stereotypes may be fathomed through investigating the linguistic patterns of the given culture. To illustrate the point, the node *south* will be investigated in the corpus of Faulkner's *Absalom, Absalom*. In particular, the corpus-derived subtext of one of the key strings containing it will be discussed.

14.3.1 The context of culture and the grammar of its lexis

According to Kozhin, culture is a set of stereotypes (Кожин 2007: 26). If a reference corpus is a balanced one and representative of a language, then the most frequent collocations in the reference corpus, whether lexical or lexico-grammatical, will reflect the cultural stereotypes of the language under discussion. It follows from this that if we are to study the context of culture of a particular epoch or social milieu, we may do so by studying the most frequent collocational patterns of its texts (viewed as language stereotypes) against the background of the most frequent patterns in the reference corpus.

The approach presented here breaks no new theoretical ground when it comes to lexical collocation. It stands to reason that frequent lexical collocates of particular nodes in the studied cultural context will differ from their most frequent counterparts in the reference corpus of a particular language. But the socio-emotional specificities (Кожин 2007: 36-37) of their use will be more fully understood if we study the grammar strings in which key lexis is embedded. Sinclair (2006, see Section 2.8) talks about the lexical collocates of 'when she was'. They are either positive (e.g. 'approached') or negative (e.g. 'raped'). Stefanowitsch (forthcoming 271-273) explores a similar approach. Both take a reference corpus as the given and study the most frequent collocations of its grammar ('she' or 'he' in this case). The approach presented here focuses on the key lexis in a particular

corpus of texts and proposes to use the reference corpus to study the grammar in which the key lexis is embedded in order to investigate semantic auras in a particular cultural setting.

14.3.2 Method

William Faulkner's major novel *Absalom, Absalom,* whose opening paragraph was discussed in Section 13.5, is one of the many he wrote describing the history of the American South. The novel starts with the description of a 'long still hot weary dead September afternoon' in the life of a southern gentlewoman Miss Coldfield and young Quentin Compson and finishes with Quentin claiming vehemently to his Northern companion Shreve that he does not hate the South. The node *south*, therefore, is legitimate for research. Below are its contexts found in the novel, in chronological order.

1 son preparing for Harvard in the **South**, the **deep South dead** since 1865 and **peopled** w
2 r Harvard in the **South**, the **deep South dead** since 1865 and **peopled** with **garrulous ou**
3 he was **born and bred** in the **deep South** the same as she was-the two **separate** Quentins
4 that there is **little left** in the **South** for **a young man**. So maybe you will enter the
5 mpson said. "Years ago we in the **South** made our **women** into **ladie**s. Then the **War** came
6 outh, since what **creature** in the **South** since 1861, **man woman nigger or mule**, had had
7face what the **future** held for the **South** but his **bare hands** and the **sword** which he at
8 parently come into town from the **south** - a **man** of about twenty-five as the town lear
9 . Yes, **fatality and curse** on the **South** and on our **family** as though because some **ance**
10 l forage wagon from Charleston, **South** Carolina and set above the faint grassy depre
11you now hold the best of the **old South** which is **dead**, and the words you read were wr
12 g them that if **every man** in the South would do as he himself was doing, would see t
13 xas or California or maybe even **South** America, **daughter doomed to spinsterhood** to l
14 or, when the very **future** of the **South** as a **place bearable** for our **women and childre**
15 as all he was after. **Jesus**, the **South** is **fine**, isn't it. It's better than the theat
16 be present on that day when the South would realise that it was now **paying the pric**
17 ot curiously at all) facing the **South** where further on in the **darkness** the pickets
18 he pickets who, watching to the **South**, could see the **flicker and gleam of the Feder**
19 sure **hate to have come from** the **South**. **Maybe I wouldn't come from the South anyway**,
20 **Maybe I wouldn't come from** the South anyway, even if I could stay there. Wait. Lis
21 blowing of the **fireflies**. "The **South**," Shreve said. "The South. **Jesus**. **No wonder y**
22 "The South," Shreve said. "The **South**. **Jesus**. No wonder you **folks all outlive yours**
23 thing more. Why do you **hate** the **South**? "I don't **hate** it," Quentin said, quickly, at

Figure 14.1 The concordance of 'south' from Faulkner's *Absalom, Absalom*!

'South' collocates with 'deep' (3), 'dead' (3), 'old' (1), 'fine' (1), 'hate' (3). There are three references to upbringing ('born and bred in the deep South'; 'hate to have come'; 'wouldn't come'), to the present state ('little left', 'paying the price'; 'you folk all outlive yourselves'), and 'future' collocates with 'bare hands and the sword' and 'place bearable' to the right of the node. There are two references to ill fate ('fatality and curse' and 'daughter doomed'. 'Family', 'ancestors', 'daughter' are mentioned once. There are several references to people

generally: 'peopled' (2), 'man' (4), 'woman' (1), 'women' (2). There is a reference to 'darkness' and there are two mentions of light of a kind that breaks up darkness ('flicker and gleam' and 'fireflies'). Overall, the picture is that of little happiness or hope for the generations of southerners.

These are the contexts containing 'Southern':

1 literary profession as so many Southern **gentlemen and gentlewomen** too are doing now
2 heless the first of the odes to Southern **soldiers** in that portfolio which when your
3 he natural thing for her or any Southern **woman, gentlewoman**. She would not have need
4 it to be. Because that's what a Southern **lady** is. Not the fact that, penniless and w
5 n the made-over dress which all Southern **women** now wore, in the carriage still but d
6 der enough since I was not only Southern **gentlewoman** but the very modest character o
7 ll, that there was actually one Southern Bayard or **Guinevere** who was no kin to y ou?

Figure 14.2 The concordance of 'southerns' from Faulkner's *Absalom, Absalom*!

Interestingly, there are two references to men ('gentlemen', 'soldiers') and six references to women ('gentlewomen', 'woman', 'gentlewoman' (2), 'lady', 'women', 'Guinevere'). Moreover, the woman is *not* of higher birth only in one case out of the six. This fact connects this concordance with context 5 of the concordance of 'south', and only this one:

"Ah," Mr Compson said. "Years ago we in the **South** made our women into ladies. Then the War came and made the ladies into ghosts. So what else can we do, being gentlemen, but listen to them being ghosts?"

Thus we arrive at an example of an investigation of a grammar string containing a key lexical item. The string 'we in the * made' yielded the following three contexts in COCA:

1 Maybe it's time ***we in the West made*** a similar bow to reality by admitting Taiwan -- the world's 13th largest trading nation (and the sixth largest trader with the U.S.) -as a full-fledged member of GATT.

2 Instead of proceeding without Aideed, the conference was held up for days until the general was finally cajoled into returning. [...] To be fair, ***we in the media made*** -- and are still making -- our own contribution to the Aideed persona. Typically, when Aideed arrived 30 minutes late to make a dramatic solo entrance into the final signing ceremony, we swallowed our disgust and ran after him for a quote.

3 No doubt we paid too little attention to potential public reaction as we ***in the industry made*** our program more international, urged on by our trading partners in the developing world. Perhaps we should have seen that the same technology that brought instantaneous trading around the world would foster instantaneous " anti " communication and organizing.

These three examples, notwithstanding the differences, all describe a situation in which someone in an inferior position should be/should not have been/was admitted to a higher status than they previously enjoyed. In two cases out of the three poorer countries join richer ones, and in one an undeserving person is elevated to the status of fame and influence by the press. As for the QPVs, 'West', 'media' and 'industry', these are all groups in possession of some influence that is exercised.

In the attempt to discover what contexts a verb in the past tense would call up in the given grammatical string, I searched 'we in the * *d' in COCA:

1 irregular reality threads (as ***we in the trade called*** them) to the omnivorous informa
2 What have ***we in the West learned*** that must be shared with the Ea
3 at he meant, at least the way ***we in the town understood*** it, was that some people had
4 d the poor, but in whose name ***we in the West carried*** out the Crusades and imperialis
5 failure to sustain attention. ***We in the UK wanted*** our troops back home almost as soo
6 to execute all three of them. We in the theatre had a sense of purpose. It wasn't ab
7 o about the fact that I think ***we in the press missed*** the voter engagement. Do you th
8 try. Then the guests left and ***we in the family visited*** for a while, and then to bed.
9 il. You're talking about what ***we in the industry called*** convergence. The convergence
10 Nor have ***we in the humanities helped*** matters; we can not even a
11 about twenty minutes of one. ***We in the congregation waited*** for him to bring up what
12, and -- absolutely not true. ***We in the board approved*** the creation of the partnersh

Figure 14.3. The concordance of 'we in the * *d' in COCA

In this concordance, seven contexts out of twelve are critical; five contexts (lines 4, 5, 7, 10, 12) criticise the action described by the verb in the '*d 'slot. To summarise, the contexts of situation above can be described as 'we, a particular group of people bound together by roots or occupation and hence sharing the same customs or expertise made a certain decision that in 41.66% of the cases we had reason later to regret'.

In the Google Books – US corpus the only expression yielded by the search line 'we in the * *d' was 'we in the West had', 'West' and 'had' being the QPVs of the string in the given reference corpus.

1 Nevertheless, ***we in the West had*** to go a long way until we were able to implement a comprehensive policy towards the East.

2 But it was all that ***we in the West had*** going for us.

3 ***We in the West had*** the masses of church members, money, education and status, and quite naturally we imposed our missional priorities and models on the rest of the world.

4 ***We in the West had*** equally unrealistic expectations about the speed of change in Russia. We tried to assist the transition, but failed to appreciate how deeply humilitating, painful and destabilising the 1990s were for the Russian people.

5 as though ***we in the West had*** nothing better to do than spy on our spiritually impoverished neighbour.

6 ***We in the West had*** to turn "nepotism" and "corruption" from tribal virtues into criminal offenses, and we struggle with it.

7 ***we in the West had*** an exaggerated notion of the repressive nature of the Soviet system.

8 we had forgotten the responsibility that ***we (in the West) had*** towards them.

9 It was bad enough that ***we in the West had*** such knowledge, but then the Soviets got it too

Figure 14.4. The concordance of 'we in the * *d' in the Google Books – US corpus

In six contexts out of nine (lines 1, 3, 4, 5, 7, 8) the West is criticised for not having shown enough understanding of the Other. This is in agreement with the previous concordance as far as the QVP 'West' is concerned. The QPVs 'West' and 'South' are semantically similar as they represent a geographical area opposed to another such area elsewhere, but this does not apply to context 5 of the concordance of 'south' from Faulkner's novel in that 'the Other' in this context is in fact the women as opposed to the men, and they all lived in the American South.

14.3.3 The context of culture in Faulkner's contexts of 'south' and 'southern'

In the studied example, the node 'south' was studied in the corpus of Faulkner's novel *Absalom, Absalom*, because not only this major novel, but also a number of his works are about the American South, and because the novel opens and closes with references to it. The concordance of 'south' shows a dark picture of no hope, whether for the past, the present or the future generations of Southerners. Additionally, the concordance of 'Southern' yielded 7 contexts from the novel, which highlight the fact that the Southern gentlewoman is one of the novel's themes. For this reason I chose to study the grammar string containing the node 'south' in Context 5 ('we in the South made our women into ladies'), having already expressed the intention of studying a grammar string in the concordance of a key lexical item. The search line 'we in the * made' in the COCA pulled

up contexts of accepting an inferior outsider, and the search line 'we in the * *d' in the same reference corpus showed that the first lexical slot was always occupied by a noun referring to a group of people sharing the same origins or expertise as opposed to possible other such groups. In addition, in the Google Books – US corpus this lexical slot was shown to be occupied by 'West', the West being a culturally specific geographical area as opposed to the Other. To conclude, the grammar string under investigation betrayed the feeling of the speaker that Southern women were actually (though deservedly) elevated to the status of men, and that they are viewed as the Other. This semantic aura has a bearing on our understanding of the cultural context of Faulkner's novel, and could not have been fully understood without recourse to reference corpora[51].

14.3.4 Conclusion

The consistency of corpus-derived subtext and its interpretive power has been proven in existing publications (Louw and Milojkovic 2016), but its role as the determinant of a particular cultural context has not been emphasized. When it comes to cultural studies in corpus linguistics, both lexis and grammar have been in the focus of research available to date (e.g. Stefanowitsch forthcoming, Masevich and Zakharov 2016), but not in the way that places corpus-derived subtext in the focus of an investigation of cultural stereotypes.

The innovative quality of this research is that it (a) starts from particular texts and looks for how these deviate from the reference corpus linguistic norm, (b) pronounces the deviation from the linguistic norm to be a secondary, and the linguistic norm to be a primary cultural scenario, and (c) looks at the most frequent lexical collocates of grammar strings as the norm, and the given lexical collocates in particular texts as the deviation from the norm that conceals the cultural norm of a particular age/milieu/text. The approach presented here does not advocate the study of grammar to the exclusion of lexical collocation and co-selection; rather, it highlights (a) the possible hierarchy of meaning creation and meaning construal processes, and (b) how such processes may be culturally pre-conditioned, when it comes to native speakers of a language. Although cultural stereotypes may differ from text to text, the language norm of a particular language remains

[51] In this case, the findings from COHA (Davies 2010-) were sparse and did not significantly change the picture yielded by COCA and Google Books US.

the frame of reference for all communication. Thus, a cultural set of stereotypes may be revealed through the grammar it employs.

Whether the context of culture be used as a frame of reference (Section 14.2) or taken as the final goal of a corpus stylistic investigation (Section 14.3), the analysis does not require deployment of any social or linguistic theory other than Contextual Prosodic.

15 Is the truthfulness of a proposition verifiable through access to reference corpora?

15.1 Introduction

The aim of this section is to explicate and illustrate the connection between corpus stylistics and philosophy of language, explored by Louw (2010a and subsequent publications). It reviews Louw's deployment of reference corpora in the light of existing philosophical and linguistic milestones when it comes to the notion of the truthfulness of a proposition. In particular, this section emphasises the link between corpus-derived subtext and the focus on insincerity in Louw's seminal 1993 paper. Louw's methodology is viewed from the point of view of Frege's Sinn/Bedeutung distinction, Russell's logical language, Wittgenstein's attitude to the relationship between language and reality, Austin's focus on ordinary language, Grice's intention behind the utterance, and later developments in the field of linguistic pragmatics. This section argues that philosophy of language with its concern for the truthfulness of propositions may be viewed as both the guarantor and the beneficiary of Louw's views, especially the work of Wittgenstein. When it comes to linguistic, and not philosophical, modality, the meanings of modal expressions in their contexts, when explored in the corpus, may be better understood. The conclusion drawn in this section is that reference corpora used inductively (Louw 2017b: 54) may shed light on the speaker's attitude to the truthfulness of their own statement.

15.2 The role of reference corpora in understanding and verifying philosophical concepts

According to the Vienna Circle, 'the truth of a proposition is the method of its verification' (Ayer 1956: 20[52], see also Louw 2013: 343). If Frege (Glock 2005: 14) held that truth is contextual (Caesar, and not Alexander, conquered Gaul), then verification of a proposition must be contextual, that is, dependent on contextual usage in the language. The phrase 'conquered Gaul' determines the truthfulness of the statement, simply because such a historical fact exists outside our personal opinion. *Sinn* and *Bedeutung* (sense and reference), as in 'the morning star' and 'the evening star' both signifying 'Venus' do not come into this so far for the same reason that the historical figures involved are in this case

[52] This source was pointed out to me by Louw.

of unquestionable identity. With meanings other than purely factual, including intentionally counterfactual ones, or those when an opinion is distinguished from what is presented as a fact, the Sense/Reference distinction grows in importance in all its contextual unambiguousness.

The verification of a proposition, or, rather, of a quasi-proposition as its manifestation in the natural language (Louw 2010a), has become possible with the advent of reference corpora for the simple reason that a language, whose balanced and representative sample is considered to be contained in the reference corpus, exists in practice through contextual usage of lexical and lexicogrammatical collocation. If philosophy of language has so far largely depended upon theory and informed intuition, reference corpora may now be considered a verification tool of philosophical hypotheses. The crucial thing to bear in mind is that all our searches must be contextually informed.

As stated in Section 2.6.2.3, Russell gives the following definition of a logical language: 'A language of that sort will be completely analytic and will show at a glance the logical structure of the facts *asserted* or *denied* […] It is a language that has only syntax and *no vocabulary whatsoever* […] *if you add a vocabulary*, [it] would be a logically perfect language' (Russell 2007: 197, emphasis added). Wittgenstein (1922) is helpful in connecting the logical structure of facts with the external world: 'The world is a totality of facts, not of things' (see Section 2.4.3), by which he invites us to consider the external counterpart to our subjective experience as a sensible whole. Going back to Russell above, human subjective experience, linguistically expressed, but in huge quantities that only computers can provide, has *logic* inherent in its *syntax*. Would *any* vocabulary make it 'logically perfect'? In the sense of 'facts asserted or denied', yes, says Russell. Since in theory any fact can be asserted or denied, then vocabulary plays a crucial part in the factuality or counterfactuality of the information given. If context decides on the distinction between *Sinn* and *Bedeutung*, then our accumulated linguistic experience is the guide that determines which *Sinn* is factual, and which is counterfactual, in relation to *Bedeutung*. Since the distinction in the quasi-propositions of the natural language (later referred to as propositions) does not presuppose only these two extremes, but the whole range in between, the nuanced (counter)factuality of our assertions or denials must also be

verified by the accumulated experience of the natural language. Russell suggests that a logically perfect language be viewed as grammar/vocabulary combinations. Lexico-grammatical collocation, then, may be viewed as what the truth of a proposition hinges upon, with the external world on the one hand, and language on the other, and the reference corpus as at the same time a sample of the language and a sample of the world, a huge collection of a particular nation's linguistic experience (Milojkovic and Louw 2017: 126). In what sense is the truthfulness of a proposition at all verifiable?

This largely depends on the notion of modality. Facts that are claimed as such must be verifiable through the external world rather than the language. Our collective linguistic experience comes in when assertions or denials are at least partly attitudinal. The truthfulness of a proposition then has to do with the genuineness of attitude rather than with what is true in the external world. Lexicogrammatical collocation in reference corpus usage must determine the speaker's attitude to facts and not facts themselves. In other words, language comes in when *Sinn* and *Bedeutung* are distinguishable.

15.3 Wittgenstein: On Certainty

If an analytic philosopher as influential as Wittgenstein, who has been informed of his own approaching death (Glock 2005: 28), chooses to dedicate the last months and days of his life to writing a philosophical work, then the themes and conclusions it contains ought to be taken most seriously. The whole of Wittgenstein's contribution to philosophy will, of course, exceed the scope of this section. He is known to change his viewpoints throughout his lifetime (Glock 2005: 11-29) to such an extent that many dissimilar notions, both linguistic and philosophical, could be defended by quoting *Tractatus* or *Philosophical Investigations*. Therefore, his last thoughts could be taken as a final comment after a lifetime of strenuous effort.

On Certainty (1969, henceforth referred to as OC), a collection of notes later formed into a book, discusses several pervading themes: the nature of knowledge as opposed to the state of believing or surmising, the extent to which a fact as provable, and the role of a system in maintaining the ideas we form of knowledge and proved facts.

Pertaining to this discussion is Wittgenstein's notion of language games. Briefly put, a language game (a) has constitutive rules, namely those of grammar, that determine what is correct and thereby define the game (language), (b) uses words whose meanings are not the objects they stand for but are determined by rules which govern their operation, that is to say, we learn meanings through usage, and (c) 'a proposition is a rule or operation in the game of language; it would be meaningless without the system of which it is a part' (Glock 2005: 193).

As Wittgenstein's work was composed over the period of 18 months, and was basically maintained as work in progress, thoughts on certain recurring themes may be grouped and cited together. Paragraphs quoted here will most often have their counterparts elsewhere in the book.

The nature of the verb 'to know' is discussed at length in many of the paragraphs, with examples taken from life, which a contemporary investigator would have searched for in a corpus. For example, 'Moore's view really comes down to this: the concept "know" is analogous to the concepts "believe", "surmise", "doubt", "be convinced" in that the statement "I know" can't be a mistake. And if that is so, then there can be an inference from such an utterance to the truth of an assertion. And here the form "I thought I knew" is being overlooked' (OC 21). Or another observation to do with the use of natural language: 'One can say "He believes it, but it isn't so", but not "he knows it, but it isn't so". Does this stem from the difference between the mental states of belief and knowledge? No' (OC 42). Or: 'When someone is trying to teach us mathematics, he will not begin by assuring us that he *knows* that a + b = b + a' (OC 113). And finally: "I know it" I say to someone else; and here there is a justification. But there is none for my belief' (OC 175). Therefore, according to *On Certainty*, the verb 'to know' is used in contexts where a certain justification is in place, but the proposition conveyed is regarded as substantiated by the speaker. To put it differently: 'It would be correct to say: "I believe…" has subjective truth; but "I know…" not' (OC 179). On the other hand, there is no knowledge that exists objectively outside the speaker: 'I act with *complete* certainty. But that certainty is my own' (OC 174). To summarise:

> ‘I know’ has a primitive meaning similar to and related to ‘I see’ (‘wissen’, ‘videre’). […] ‘I know’ is supposed to express a relation, not between me and the sense of a proposition (like ‘I believe’), but between me and a fact. So that the fact is taken into my consciousness. (Here is the reason why one wants to say that nothing that goes in the outer world is really known, but only what happens in the domain of what are called sense-data) (OC 90).

Wittgenstein’s resorting to natural language when defining the scope of the verb ‘to know’ is supported by his conviction that meanings of words are contextually defined, which is another link that connects *On Certainty* to the reference corpus as a source of real-life contexts of situation: ‘A meaning of a word is a kind of employment of it. For it is what we learn when the word is incorporated into our language. That is why there exists a correspondence between the concepts “rule” and “meaning” (OC 61-62). Wittgenstein compares the meaning of a word with the function of an official, in the sense that an official can have different functions (OC 64). Whole propositions are contextually dependent: ‘the spoken or written sense “2x2=4” in Chinese might have a different meaning or be out and out nonsense, and from this is seen that it is only in use that the proposition has a sense’ (OC 10).

This idea, later incorporated by Firth in his notion of ‘meaning by collocation’ (Firth 1957: 181) leads us to the notion of language games: ‘If we imagine the facts otherwise than as they are, certain language games lose some of their importance, while others become important. And in this way there is an alteration – a gradual one – in the use of the vocabulary of a language’ (OC 63). Since we have seen that language games are based on the rules of grammar, this view may be construed as echoing Russell’s ‘logical language’, with grammar acting as the game-changer when the facts are not presented as they are – in Russellian terms, their logical structure being ‘asserted or denied’. In terms of the present work, the vocabulary discussed is mainly ‘know’ and verbs such as ‘believe’. The representation of facts changes, rules change, vocabulary changes also, albeit ‘gradually’. The ‘true’ description of facts, whatever they are, seems to be connected to changing rules of the language and changing vocabulary. The different – and extremely

nuanced – semantic auras shown by the reference corpus, such as those discussed in Section 15.4 would have satisfied Wittgenstein.

> Our knowledge is rooted in a system of beliefs:
>
> 141. When we first begin to *believe* anything, what we believe is not a single proposition, it is a whole system of propositions. (Light dawns gradually over the whole.)
>
> 142. It is not single axioms that strike me as obvious, it is a system in which consequences and premises give one another mutual support.
>
> 144.[…] What stands fast does so, not because it is intrinsically obvious or convincing; it is rather held fast by what lies around it.

Mistakes are possible only in relation to the system: 'In order to make a mistake, a man must already judge in conformity with mankind' (OC 156). Our system of beliefs existed prior to our existence: 'But I did not get my picture of the world by satisfying myself of its correctness; nor do I have it because I am satisfied of its correctness. No: it is the inherited background against which I distinguish between true and false' (OC 94). This is not stated as reason to doubt the system, but to describe the nature of our knowledge of the external world, our 'world-picture' (OC 95). It is a logical system (OC 136). It is not about experience, according to Wittgenstein: 'Experience is not the ground for our game of judging. Nor is it an outstanding success' (OC 131). If a person claimed it was possible to have been on the moon, '[w]e should feel ourselves intellectually very distant from someone who said this' (OC 108), despite our inability to prove though experience that it is impossible.

What change would Wittgenstein's ideas have undergone had he been brought in touch with reference corpora? He abided by the view that all philosophical problems were rooted in language (Glock 2005: 236). Our beliefs, and even doubts, form a system (OC 126). These must be rooted in language also, and language creates meanings through use. Meanings must be the entity that mirrors the system of our beliefs, or the knowledge of learned, acquired facts. According to Wittgenstein, if we present facts differently from what

they are, the rules of the language game change, and these encompass grammar and vocabulary. How would he have gone about extracting these changes from the corpus?

Ideally, he would have done it through calculation. In the whole text of *On Certainty*, it is calculation that he trusts most. 'So is the *hypothesis* possible that all the things around us don't exist? Would that not be like the hypothesis of our having miscalculated all our calculations?' he asks (OC 56). It is not possible to miscalculate *all* calculations. And calculations are precious because they are concerned with experience, not rules:

> 44. If you demand a rule from which it follows that there can't have been a miscalculation here, the answer is that we did not learn this through a rule, but by learning to calculate.
>
> 45. We got to know the *nature* of calculating by learning to calculate.

Calculation is much better than certainty: 'When someone has made sure of something, he says: "Yes, the calculation is right", but he did not infer this from his condition of certainty. One does not infer how things are from one's own certainty' (OC 30). Reference corpora allow calculation in the sense that the language norm in the form of the most frequent collocations may be objectively deployed in calculating meanings in particular contexts. This is finding out through a process, not a rule. For example, Wittgenstein introduces the idea of normal circumstances: 'If, however, one wanted to give something like a rule here, then it would contain an expression "in normal circumstances". And we recognise normal circumstances but cannot precisely describe them. At most, we can describe a range of abnormal ones' (OC 27). With the advent of reference corpora, such conditions have become describable. Ideally, interpretation has been reduced to a minimum. 'At some point one must pass from explanation to mere description' (OC 189).

15.4 Wittgenstein and Brodsky, or how much Brodsky knew about language games

As previously stated, if Wittgenstein had been informed of his own upcoming demise, and was spending his last days on a piece of writing, he must have thought its themes meaningful to his own philosophical thought in particular and to philosophy in general. Iosif Brodsky, arguably the most famous dissident poet of the Soviet era, was

concerned with death all his life, as he suffered from a weak heart. All these background details are seemingly of no interest to a linguist, but they matter because of similar contexts of situation (Firth 1957: 182). At least, when it came to philosophy, in the case of the former, and death in the case of the latter, both these men knew what they were talking about.

Brodsky's poem *Он знал, что эта боль в плече...* (*He knew that this pain in the shoulder..*, 1965) starts with the persona feeling his regular cardiac pain, of which he, however, dies in the last stanza (see Section 7.5). What concerns us in the poem is the use of modal expressions which, as it will later be shown, grows into a poetic device. When summed up, their ostensible use indicates that during the first half of the poem the persona refuses to believe he is going to die. These are the expressions in their contexts:

Он знал, что эта боль в плече уймется к вечеру…
(He knew that this pain in the shoulder would abate by the evening…)

Он вообразил,что боль способна обмануть…
(He imagined that pain was able to deceive…)

…он, считавший: ежели сполна что вытерпел - снесет и впредь…
(…he who considered that if he had withstood something before, he could do it again…)

…не мог представить, что она его заставит умереть.
(…could not imagine that it would make him die.)

Он знал, что ('he knew that') was searched in the Russian National Corpus (later referred to as the RNC). In the first fifty contexts, ten had to do with life and death situations. In all of them, the search line is followed by a reference to death; in nine of these ten, it is the agent designated as 'he' who is going to die; in eight of them, the agent is the one who knows he is going to die. The rest of the fifty contexts describe situations other than life and death ones.

Он вообразил, что ('he imagined that') yields 47 contexts in the RNC. Two contexts are an exception: what the agent imagined later indeed turns out to be the case, but

in both situations there are no clues in reality pointing to the possibility of such an outcome. In the rest of the contexts, characteristically, what the agent imagines is ridiculous, laughably impossible. Furthermore, in seven contexts out of 47 the agent who is doing the imagining is actually a madman.

Он считал, если ('he considered that') is a slight and more frequent modification of the actual form used by Brodsky but absent in the RNC. This expression yielded only four contexts. All these contexts suggest that what the agent considers to be true (and this is always a rule of thumb of some sort) is an idealistic notion of life and is not rooted in reality.

Не мог представить, что ('could not imagine that') in Russian starts with the verb phrase but retains the grammatical masculine gender. If the available data is narrowed down only to those contexts where the future tense obtains to the right of the node, the corpus yields three categories of contexts: (a) in eight contexts the agent could not imagine that such a beautiful thing would happen (which later really happened); (b) in nine contexts the agent could not imagine that such a horrible thing would happen (which later did indeed happen); and (c) in five contexts the actual English translation should be 'could not imagine what' (e.g. he would say at the meeting), but perhaps it is still worth stating that the contexts of situation are negative in four cases out of the five. The context in Brodsky falls under the second category.

In the first half of the poem the author clearly uses modal expressions as device. On the conscious level, he uses them to build the following gradation:

He knew that he would not die…

He imagined that it was possible for him to die…

He considered that he was able to withstand the pain as before…

He could not imagine a contrary scenario.

Authorial intention is complemented by the reference corpus, in which each of these modal expressions is surrounded by an additional semantic aura (provided we as researchers have ensured the similarity of the context of situation during our investigation):

He knew that (if the context is one of life and death, the agent dies)…

He imagined (what is utterly improbable)…

He maintained an idealised view of reality…

What he could not imagine would actually take place.

These additional semantic auras are contextually inherent in the language norm and add to the pathos of the hero's tragic demise. From the purely linguistic point of view, what matters to the researcher is not what the persona knows, but what Brodsky knows. The persona could not have known (shall we say) that he was going to die, but Brodsky could and did. There are two kinds of contradicting the ostensible meaning of the modal expressions discussed: in reality, the persona dies; in the language, the corpus contradicts the persona throughout (except in the case of 'he imagined that', where the corpus tends to aggravate the hypothetical nature of the statement rather than diminish it). This consistency forms a Wittgensteinian *system*.

These four expressions are not the same in implication, because Brodsky starts off using the verb 'to know'. It is, according to Wittgenstein, different from all the other expressions used in the poem ('to imagine' etc.) because it is less subjective, carrying a message that is supposed to be able to involve another and still hold valid. The beginning would have been far less effective had Brodsky used a more tentative modal expression. This is the nature of his language game: 'knew' becomes 'imagined', 'that' turns into 'if', then negation sets in. First vocabulary, and then grammar are brought into play. To his credit, Wittgenstein did not have a corpus.

In Wittgenstein's terms, the system of beliefs we live by is even more deeply ingrained in our use of language than he knew it to be. Better still, it is more deeply ingrained in it than what we ostensibly consider our system of beliefs. Obviously, there emerge two layers of modality: the hedge towards the world that we erect by using modal expressions (*Sinn*), and the actual meaning of these expressions as it emerges when studied in the corpus in as similar contexts as possible (*Bedeutung*). *Bedeutung* here is not only strictly contextual, but also based on the process of *calculation*, so favoured by Wittgenstein, and not a rule that we are taught. Also, generalising on particular expressions such as the verb 'to know' does not seem to be enough: they ought to be studied (a) as part of their collocational patterns, and (b) in their contexts of situation. Distinctions such as

how 'to know' differs from 'to believe' or 'to imagine' are only a starting point. If meaning is use, the role of the reference corpora is to provide its precise definition.

In Wittgenstein's terms, the persona imagines the facts otherwise than as they are, and this brings about the gradual alteration of his vocabulary and then grammar (rules of the game). In Russell's terms, the logical structure of the facts is being denied. The rules of Brodsky's language game are such that the reference corpus recovers the logical structure of his facts. Going back to Wittgenstein, our system of beliefs is inherent in the language, not in our system of beliefs.

15.5 Shared logical form

The question of the truth of a proposition in the context of collocation was first brought up by Louw (1993) – in those days, still without the mention of philosophy. Louw states that if there is collocational mismatch between authorial text and the reference corpus norm – in those days, still at the level of the lexical – then either the speaker is being ironic, or the speaker is being insincere. The idea of irony has grown into the project (still only tentatively attempted, see Section 7) to give all literary devices their corpus definitions, while the idea of insincerity has been defended in a number of publications (as a means of characterization, or evidence of lack of inspiration), mainly against too literal interpretations (see sections 3.2.4 and 3.2.7). Philosophy of language started to play a role in Louw (2010a), with corpus-derived subtext (as the most frequent lexical collocates of grammar strings, termed, as was previously stated, quasi-propositional variables) explained and exemplified in Louw (2010b). The idea of the Vienna Circle that the truth of a proposition is the method of its verification has led him to claim that collocation is instrumentation for meaning. Going back to Louw (1993), if the speaker violates the language norm that puts him straight, he cannot have been straight at least with himself.

This notion has to do with shared logical form. 'The logical form of a proposition is its structure as paraphrased by formal logic for the purpose of revealing those features which matter to the validity of arguments in which it occurs' (Glock 2005: 212). Dating back to Aristotle, the idea was revived by Frege's function-theoretic analysis, and taken up by Russell, who understood philosophy as analysing the logical form of propositions

(Glock 2005: 212). Wittgenstein viewed a proposition as a picture 'which models reality, *truly or falsely*, by virtue of the relationship between its elements representing the relationship between the elements of a situation. Such a picture possesses two essential features, firstly a method of projection connecting *the elements of the model with the elements of the situation it represents*, and secondly *structural features which it must share with reality in order to depict the latter*… "Logical form" is whatever picture… must share with what it depicts' (Glock 2005: 213; my italics).

Given that reference corpora contain accumulated language experience that reflects accumulated understanding of the world, shared logical form must reflect shared experience: 'Just as each proposition must share its logical form with the state of affairs it depicts, so language, the totality of propositions, must share with what it depicts *the* logical form, "the form of reality" (TLP2.18 – apparently equivalent to the "form of world")' (Glock 2005: 215).

Russell's logical language is founded on syntax, which becomes the core of experience, with vocabulary added to perfect what Wittgenstein called the 'pictorial form' (Glock 2005: 213). Where the facts are described as 'imagined otherwise than as they are' (OC 63), the rules of the language game change, and falsity (the 'imagined' quality) is detected through lexis and then grammar, as in Brodsky. As for the interconnectedness of the semantic auras in the poem by Brodsky (see also Section 7), it must be ascribed to the logical relations between propositions (Glock 2005: 331).

A proposition, then, may mirror what is true or state what is imagined. The ultimate task is to differentiate between them. What is true will correspond to the logical structure of reality. Its logical form, then, will be identical to other logical forms in the corpus in a particular state of affairs mirroring reality. If the logical form is identical, then the variables will change the 'picture of the world'. According to Wittgenstein, propositions 'can depict reality either truly or falsely only by being pictures or models which are compared with reality, as a ruler is laid next to an object of measurement […] A picture agrees with reality or fails to agree; it is correct or incorrect, true or false […]' (Glock 2005: 366).

Austin (1950), however, rejects Wittgenstein's picture theory. He starts with qualifying statements featuring the adjective 'true' and insisting on the preposition 'of' (not

'to', as in 'this picture is true to life'). Apart from not recognising truth as residing in pictures (Austin 1950: 112), he departs from Wittgenstein in that he focuses more on the nuances of ordinary language. The true/false dichotomy turns into a scale: 'We say, for example, that a certain statement is exaggerated or vague or bold, a description somewhat rough or misleading or not very good, an account rather general or too concise. In cases like these it is pointless to insist on deciding in simple terms whether the statement is "true or false" […] There are various *degrees and dimensions* of success in making statements: the statements fit the facts always more or less loosely, in different ways on different occasions for different intents and purposes' (ibid.: 124). The importance of context is emphasised: '[T]he question of truth and falsehood does not turn only on what a sentence *is*, nor yet on what it *means*, but on, speaking very broadly, the circumstances in which it is uttered. Sentences *as such* are not either true or false' (Austin (1962), quoted from *The Stanford Encyclopedia of Philosophy*).

Austin dwells on the distinction between a proposition and its meaning: 'we never say: "The meaning (or sense) of this sentence (or words) is true: what we do say is what the judge or jury says, namely that: "The words taken in this sense, or if we assign to them such and such a meaning, or so interpreted or understood, are true" […] here "words" and "sentence" refer […] to the words or sentence *as used by a certain person on a certain occasion.* That is, they refer […] to *statements*' (Austin 1950: 113). Sentences cannot be true or false, statements can. It is misleading to conclude that a statement is true when it corresponds to facts (ibid.: 115). In fact, a statement may be said to be true if it respects demonstrative conventions (correctly refers to 'historic' situations, that is, to the world) and descriptive conventions (uses words of the type that is used to describe such a historic state of affairs) (ibid.: 116). The trouble with the word 'facts', according to Austin, is its linguistic realization: phrases like 'it is a fact that' make the word 'fact' synonymous with 'true statement' and nothing else (ibid.: 117). Austin writes: "Fact that" is a phrase designed for use in situations where the distinction between a true statement and the state of affairs about which it is the truth is neglected […] speaking about "the fact that" is a compendious way of speaking about a situation involving both words and the world […] the correlation between the words (=sentences) and the type of situation, event, etc. which

is to be such that when a statement in those words is made with reference to a historic situation of that type the statement is then true, is *absolutely and purely* conventional' (ibid.: 118). While Austin thus retains Wittgenstein's link between linguistic conventions and the world as an objective entity, he departs from Wittgenstein significantly in rejecting a direct connection between the structure of the statement and the structure of the world: 'there is no need whatsoever for the words used in making a true statement to "mirror" in any way, however indirect, any feature whatsoever of the situation or event; a statement no more needs, in order to be true, to reproduce the "multiplicity", say, or the "structure", or "form" of the reality, than a word needs to be echoing or writing pictographic. To suppose that it does, is to fall once again into the error of reading back into the world the features of language' (ibid.: 119).

By introducing demonstrative and descriptive conventions, Austin does everything in his power to bring together the world and the language it describes. There is nothing wrong either with his scrupulous adherence to natural language. Austin simply refuses to take Wittgenstein literally. Still, in the absence of reference corpora which can be said to subsume both conventions introduced by Austin, he chooses to study meaning separately from a sentence. The language data found in a reference corpus would have given enough empirical material to correlate demonstrative and descriptive conventions on a large scale, and describe what Wittgenstein refers to as 'normal circumstances'.

Grice (1957) also connects meaning to facts. To him, a sentence like 'Those spots meant measles' may be reformulated as 'The fact that he had those spots meant that he had measles' (Grice 1957: 377). This is, to him, meaning in the natural sense (as are expressions such as 'X means to do…'). But if he uses the verb 'to mean' in the sense 'to explicate', 'to intend to convey', that is the word's nonnatural sense (ibid.: 378-9). A causal type of answer to the question 'What is meaning?' is rejected because inference in situational contexts does not entail linguistic rules. If we see a person wearing a tailcoat, we may assume they are going to a dance; if a person is an athlete we may assume they are tall. In both cases we might be mistaken. Inferences from utterances may be just as likely, but not obligatory. Also, general meanings of conventional signs 'make no provision for dealing with statements about what a particular speaker or writer means by a sign on a

particular occasion (which may well diverge from the standard meaning of the sign); nor is it obvious how the theory could be adapted to make such provision' (ibid.: 380-381). Thus Grice introduces the distinction 'on particular occasion vs. timeless', within nonnatural meaning, and then defines a true statement with the help of another distinction: intention. A statement 'x means something' is true if the utterer intends to create a belief in the audience. The belief is to say what x means. This is not sufficient: the utterer must have had 'intended the audience to recognise the intention behind the utterance' (ibid.: 381). Truth now involves not only intending to create a belief, but intending to create a belief about the intention behind the utterance. Nonnatural meaning, according to Grice, subsumes cases in which there is recognizable intention, e.g. frowning without nonnatural meaning vs. frowning in order to convey displeasure (nonnatural meaning present). But to the nonnatural meaning of the utterance itself only the primary intention is relevant, e.g. giving information, not the motive with which the information is given (ibid.: 382-386).

What is formulated by Grice gets very close to what is formulated in Louw (1993) and his research to date: that authorial intention (see also Section 9 for the connection between inspired text and insincerity and Section 8 for the description of manipulative discourse) is traceable by comparing what is the language norm to what deviates from it:

> Explicitly formulated linguistic (or quasi-linguistic) intentions are no doubt comparatively rare. In their absence we would seem to rely on very much the same kinds of criteria as we do in the case of nonlinguistic intentions where there is a general usage. An utterer is held to intend to convey what is normally conveyed (or normally intended to be conveyed), and we require a good reason for accepting that a particular use diverges from the general usage (e.g. he never knew or had forgotten the general usage). Similarly in nonlinguistic cases: we are presumed to intend the normal consequences of our actions. Again, in cases where there is doubt, say, about which of two or more things an utterer intends to convey, we tend to refer to the context (linguistic or otherwise) of the utterance and ask which of the alternatives would be relevant to other things he is saying or doing… (Grice 1957: 387)

Thus, seven years after the death of Wittgenstein, the definition of truth as meaning conveyed whose intention corresponds to what the audience believes is its intention enters the scene and expands Frege's Sinn/Bedeutung distinction. Logical form is ignored in favour of natural language in context, which may yet be remedied through recourse to reference corpora. Crucially, Grice finishes by saying that 'linguistic intentions are very like nonlinguistic intentions' (Grice 1957: 388).

The later developments in the field of pragmatics, despite their focus on context and natural language, contribute little to the topic and argument of this section. The speech acts theory, developed by Austin and Searle (Douthwaite 2000: 68-74), does not assist in explicating nuanced subliminal semantic auras that are conveyed to the addressee as part of the whole utterance. Clearly, in their desire to adhere to context, pragmaticists retain the Sinn/Bedeutung distinction, but abandon logical form, because in the absence of the reference corpora it is their precomputational best. Louw's Contextual Prosodic Theory seems able to resolve these difficulties. His notion of *shared logical form* (Louw and Milojkovic 2015: 542-544) resides in the primacy of grammar over vocabulary, the former, however, adding to the final construction of meaning through departure from the most frequent collocations in nuanced context-bound directions determined by *repeatable events*. This process is not mechanical because it is not read into the language by the analyst, but exists independently even of the speakers' intuition. It simply is. If the interaction between the logical form and the vocabulary it employs contradicts the ostensible meaning conveyed by the speaker, the truth-value of the proposition is at stake. As was claimed by Grice (1957: 387), the intention to convey untruth can only be pinpointed by context. From this work onwards, pragmaticists become prisoners of context and the potential multiplicity of meaning to the exclusion of contextually bound code, which, through the reference corpus, may both explicate meaning and comment on the truthfulness of propositions: 'Where *p=non-p*, illogical propositions will be identified by logic and must then be shown the door' (Pears: 1971: 70, quoted in Louw and Milojkovic 2016: 150).

15.6 Conclusion: prospection as a form of insincerity

The poem by Brodsky, analysed in the previous sections, is a perfect example of *prospection* identified by corpus-derived subtext: the meaning ostensibly expressed in the first line giving way in the reference corpus to the key motifs explored in the subsequent text straight to its conclusion: 'He knew that he would not die' changes to 'he knew that he would die' in the corpus (for discussions of the mechanism of prospection at work in English poetry, see Louw and Milojkovic 2015, 2016). Despite the differences related to genre, length being the most obvious one, the opening lines of novels also show prospection, as was demonstrated at length in Section 13. In the first paragraph of DeLillo's *White Noise*, the first grammar string '*s *ed at' not only supports the motif of looking (the theme of the passage), but also that of higher education and research (the theme of the novel) and even armed conflict (the ending to the novel). Both authors somehow seem to incorporate the ending of their text, long or short, into their opening line. The truth of their proposition is at stake and revealed by the reference corpus through logical form: syntax. Logical form appears to act as a means of building fictional worlds.

The overview of philosophical and linguistic developments in the field of the truthfulness of a proposition, attempted in this section, shows that the methodology suggested by Louw (1993 and subsequent publications) is best vindicated by Wittgenstein's adherence to logical form. Wittgenstein, unlike Austin and Grice, believed in the correlation between logical form expressed by language on the one hand, and the structure of the world on the other. Later philosophical developments, which led to the growth of pragmatic theories, questioned this direct connection and may be said to have focused on context to the detriment of exploring logical form. Wittgenstein's (and Russell's) ideas can now be revived through studying the nuanced relationship between syntax and vocabulary in the reference corpora, contextually explored.

It appears that any, just any combination of words used in a certain context in the natural language will have its contextual semantic aura in the corpus. These may be modal expressions, and these can be unqualified propositions. The only condition is that the expression be studied in its contextualized linguistic manifestation.

Perhaps it could roughly be stated that any lexico-grammatical collocation, used in agreement with the language norm (the most frequent collocations in the reference corpus in their contextual realization) supports the speaker's utterance in the sense that his view of his own proposition is objective. If, however, there is a mismatch between the aura found in the corpus and the ostensible meaning of the proposition as is, for example, found in Brodsky, the objectivity of the speaker with regard to his/her own proposition is at stake and the possible world that his utterance has created has conflicted with, or at least deviated from, the possible world created by the sufficient corpus data. If understood this way, shared logical form (wildcarded vocabulary which is inaccessible to intuition) becomes part of alethic necessity. Thus, in view of the depths that it plumbs, corpus linguistics attains the status of a new theory.

This would enable us to deal not only with expressions that are traditionally viewed as modal, but also with all propositional content that hides additional meaning. It would be the case because what the speaker believes to be the illocutionary force of his statement would be subject to revision, even though the revision was initiated not by the speaker or the addressee, but by the researcher (and in Turing's case, by the machine, see Louw 2017a).

Thus, the reference corpora reveal the speaker's attitude to their own utterance, albeit beyond the speaker's knowledge or intention. The speaker evaluates both their facts and their proposition from *within* the proposition itself. The difference between the ostensible and superficial and the implied and profound is the measure of the speaker's attitude to the truthfulness of his or her proposition. Both Brodsky and DeLillo intend to write about death, and this is shown in line 1. This type of *insincerity* (Louw 1993) has little to do with crude attempts at deception (see Section 8): research has shown that in such cases the speaker uses frequent lexico-grammatical combinations which in the reference corpus call up a coherent class of completely different states of affairs, not sharing situational aspects with the context of the utterance but instead consistently skewing the picture in the direction desired by the speaker.

In the same way in which Wittgenstein was optimistic as to the nature of knowledge, Russell expressed optimism as regards the existence of truth: 'I conclude that

"truth" is the fundamental concept, and that "knowledge" must be defined in terms of truth… This entails the consequence that a proposition may be true although we can see no way of obtaining evidence either for or against it…' When it came to knowledge, he settled for 'principles of *non-demonstrative inference* which may be difficult to reconcile with *pure empiricism*' (Russell 1996: 22; my italics). It is to be hoped that, when it comes to the knowledge of whether the speaker agrees with his or her own proposition or, which is the same thing, whether he or she considers it truthful, the pure empiricism of reference corpora, if cautiously used, may prove of some value. A different point, however, is worth raising here: if Brodsky and DeLillo know the fate of their protagonists and say so, may we study our own speech in the hope of predicting, and then even influencing, our future?

16 Discussion

Contextual Prosodic Theory (CPT), developed by Bill Louw (starting from 2000), is a corpus stylistics theory which compares authorial collocation in context to that same collocation contextualised in the reference corpora. It has been used for a variety of purposes, which can be described as semantic (revealing the nuances of meaning), stylistic (offering corpus-attested definitions of literary devices), those of translation (adjudicating the semantic equivalence of the translated text), and pedagogical (teaching all these techniques within the framework of higher education). As for nuances of meaning, particular attention has been paid to *prospection*, the ability of the first line of a text to contain grammatical subtext (the semantic aura of the most frequent lexical collocates of a grammar string) which can be derived by means of a corpus and which points to themes that are to follow.

Contextual Prosodic Theory originated in the broader academic context of Sinclairan (neo-Firthian) corpus lingustics and adheres to its main principles (reliance on raw data, bottom-up analysis, collocation-via-concordance rather than automated statistical collocation, and detailed study of context). Louw's published works to date are mostly of theoretical nature and contain relatively few practical examples. His thinking has evolved beyond literary criticism and entered the domain of philosophy of language, laying particular emphasis on Russel's logical language as the foundation of corpus-derived subtext and on the notion of proximity as entailment and of consilience of induction (combining corpus linguistics, stylistics and philosophy of language). As for commentaries by other academics, they have mostly commented on corpus-derived subtext in the domain of stylistics and on semantic prosody in the area of corpus linguistics. Generally speaking, when it comes to the influence of CPT on corpus stylistics as a discipline, it mostly makes use of semantic prosody and generally relies on corpus linguistic methods of studying recurring patterns in (literary) texts. Corpus-derived subtext has not yet gained as much popularity as semantic prosody as a method of analysing meaning in texts. This should change in the near future because evidence is emerging that meaning is not individually constructed, but shareable between readers. This is confirmed by several psychological

experiments supporting the impact of semantic prosodies on meaning construal, as well as by the phenomenon of prospection as evidence that semantic auras in a text are interrelated. These point to the conclusion that semantic auras researched by CPT are an aspect of general semantics rather than offering a possible explanation for individual interpretations.

However, the study of a wording in a famous poem by Pushkin shows that authorial lexical usage may differ from the one generally present in the reference corpus, highlighting the need to consult the author's corpus alongside the general reference one. The reason for the mismatch in this case is the diachronic change that the SP of the lexeme 'war' appears to have undergone in the Russian language. However, a similar mismatch may be the result of an author's particular stylistic preference, therefore such cases may have to be explained by reasons other than diachronic changes. This is shown by the analysis of Pushkin's other combination, this time lexico-grammatical. The analysis of frequent lexical collocates of its grammar (corpus-derived subtext) shows a subtle negative meaning not accessible to the native speaker of Russian, but confirmed by the diachronic Poetry corpus of the RNC. This delineates the distinction between authorial intention and textual meaning, and points to the ability CPT to diagnose 'authorial insincerity': an instance of an author saying more than he or she may have intended. This can be done because the semantic auras of grammar strings are completely opaque to intuition. The second segment in question also assists in drawing conclusions as to the degree of originality of rhetorical devices, which can be established based on the relexicalisers present in the text: the lexical meaning may be relexicalised or delexicalised, depending on co-occurring lexemes. This part of the analysis compares Pushkin's poetry (in Russian) to that of W. B. Yeats and Herbert Williams (in English). Overall, one of the aims of studying Pushkin's wording was to confirm the theory's transferability to Russian.

The ability of CPT to comment on both Russian and English semantic prosodies and corpus-derived subtext offers opportunities to analyse existing Russian-English translations and produce new ones. This, of course, does not preclude the researcher from applying similar strategies to English-Russian translation, or contexts of Slavic languages and English. The analysis of an existing translation shows that semantic auras of lexis and grammar may not be adequately transferred to the target language, to the point of creating a

completely opposite implication when it comes to grammar. It follows from this analysis that states of affairs created both by lexical and grammatical collocations in the source language must be carefully checked in the reference corpus of the source language, and then the same procedure must be applied to the translation. This does not look practicable because of the necessary scope of the analysis; however, two analyses of professional translations performed in Section 6 confirm that a talented native speaker's intuition helps to render an acceptable equivalent in the target language. It is recommended that translation be performed by both a talented professional and a native speaker of the target language, who will occasionally check lexical and grammatical auras in both reference corpora, particularly at junctures where he or she experiences difficulties. Also, a translator may consult specific contexts in the reference corpora and from them gain experience of a particular emotion that must be conveyed to the translation.

The analysis of Pushkin's wordings and of the ways in which they may be translated focuses on isolated textual segments and does not include a study of interaction of corpus-derived subtext within a text. Such an interaction is studied in W. B. Yeats's poem 'The Circus Animal's Desertion'. Since the analysis of this poem reveals that it contains two metaphors with the same referent and very similar subtext, it then arrives at a corpus-attested definition of metaphor (in the spirit of Louw 2008) and compares it to a corpus-attested definition of pun. Also, on the basis of a poem by Brodsky, a corpus-attested definition of personification is given. Although corpus-attested definitions of figures of speech have an advantage over intuitive definitions in that they are more empirical, the analysis offered is programmatic in nature: it shows how to answer a particular research question rather than take up an issue and resolve it fully. The corpus-attested definitions of devices presented here are a by-product of analysing semantic auras of grammar strings within a particular poem. Although the semantic auras of grammar strings are completely opaque to intuition, their interaction within the poem is shown to be consistent and coherent.

Formulating definitions of rhetorical devices answers one aspect of the irony/insincerity dichotomy postulated in Louw's seminal paper 'Irony in the Text or Insincerity in the Writer: the Diagnostic Potential of Semantic Prosodies' (1993). The other

aspect is insincerity. Authorial insincerity on Pushkin's part, found in one of his lexico-grammatical collocations, is not manipulative. On the other hand, manipulative discourse is not found to contain miscollocations: rather, it contains a far greater proportion of frequent lexico-grammatical collocations belonging to a context of situation that the liar wishes to construct in the minds of the audience. These results are obtained by tackling Russian newspaper discourse already diagnosed as manipulative by a group of Russian forensic experts.

Findings so far discussed enable the construction of a cline of prosodic and subtextual clashes, aiming, ideally, to explain all discourse from the point of view of CPT. The cline starts from cliché and banality, then examines sincerity, 'inspired sincerity' containing lexical mismatches, 'inspired insincerity' denoting subtextual clashes, and 'uninspired insincerity', giving corpus-attested definitions of each kind of deviation from the norm as found in the reference corpus. It is assumed that cliché, banality (not catering to the needs of the context when it comes to the expression of an emotion) and uninspired insincerity do not belong to quality discourse. In contrast to the division proposed by Milojkovic in Louw and Milojkovic (2016), not all 'uninspired insincerity' is equated with lying. Forced discourse (such as uninspired essays, white lies and even inept translations) occupy the space after 'inspired insincerity'. The next slot is taken up by manipulative discourse (intentional lying), which seems to grow into cliché and to turn the cline into a circle. This division is also an attempt at proving that prosodic and subtextual clashes which have intentional or unintentional artistic effects are the consequence of the state of mind known as artistic inspiration. Successful literary devices would then occupy the slot between inspired sincerity and inspired insincerity, because they involve both collocational and subtextual clashes. The notion of inspiration is also explored as it is understood by a particular British poet, opening several unresolved issues but confirming that prosodic and subtextual clashes may take place moments of inspiration.

The cline of prosodic and subtextual clashes does not take into account the phenomenon of prospection, which should be given a slot under the category of 'inspired insincerity', because it is a property of corpus-derived subtext. In two different novels, by Faulner and by DeLillo, prospection is found to operate in the very first line or passage, to

cover some of the major themes and motifs of the novel, and, in the case of DeLillo, to anticipate the very ending of the text. Therefore, prospection so far seems the most reliable proof of the interrelatedness of semantic auras of within a text.

Since corpus-derived subtext generally plays a part in understanding textual meaning, it can show how textual meanings may combine to build a picture of a particular text's cultural milieu. For example, an alalysis of a line from Faulkner's *Absalom, Absalom* shows that women in the American South may have been perceived by men as having an inferior status.

Two experiments were conducted with university students who were offered an opportunity to explore texts in accordance with CPT principles. While the emphasis in the first experiment was on lexical collocation and the semantic auras of grammar strings, the second focused on corpus-derived subtext. The results of the first study were encouraging, because the distribution of the marks correlated with what could generally be expected of a university course. The second study, dealing with subtext, was additionally useful in highlighting stumbling blocks that could be encountered when teaching CPT.

Generally, corpus stylistics is understood by Louw to have philosophical foundations. Russell's understanding of logical language in fact corresponds to corpus-derived subtext in natural language. Russell's 'facts asserted or denied' are meanings created by lexical slots in grammar strings being filled by expected or less expected lexical collocates. The expected ones constitute the subtext of the string (facts asserted), and the unexpected ones are facts denied, the picture of the world altered either as struggling insight at moments of inspiration, or as a figure of speech, or as forced expression. In this way, grammar is seen as the logical form shared by all of the expressions of which it is a part, pointing up insincerity (a skewed picture of the world) by any odd lexis that has been detected in place of what is expected or what is creative. This seems to be what Wittgenstein called 'language games', which are explained by using Brodsky's poem *'Он знал, что эта боль в плече...'* as an illustration.

All in all, Louw's inventions, relexicalisation and corpus-derived subtext, appear to work similarly both in English and Russian. The mechanism of subtext containing additional layers of meaning has been illustrated by its working across the whole span of

poems chosen - 'The Circus Animals' Desertion' by W.B. Yeats and *'Он знал, что эта боль в плече...'* by Iosif Brodsky. The objective role of subtext in meaning creation is underscored by prospection discovered not only in short poems by celebrated poets, but also in Serbian students' English essays and in novels by Faulkner and DeLillo. The interaction between corpus-derived subtext and meaning is evident and cannot be put down to coincidence.

An important aspect of discussing and presenting Louw's theory is the element of the unexpected that verges on the surreal. It seems implausible that two metaphors with the same referent but in opposite sections of a poem (its opening stanza and its coda) should have such similar groups of QPVs as their subtext. Neither could it be predicted that Brodsky's beginning prospects death in the corpus and in the poem, but is optimistically used by the unsuspecting - or suspecting? - persona. The negative subtext in the greatest love poems in the Russian language by Alexander Pushkin will come as a shock to many Russian readers - here Pushkin's biographical data proved most useful alongside corpus findings.

The research presented here is not without its limitations. The first is related to the nature of procedures that need to be carried out in order to establish the meaning of a stretch of text by CPT methods. Establishing semantic auras in the reference corpus is a time-consuming task. This is certainly worth doing in order to ascertain the meaning of a famous poem, or to hand in a forensic report. However, the time taken by consulting numerous contexts might seem inordinately long. From this follows another limitation: some analyses, particularly to do with Pushkin, needed to be presented on pages of text. This has led to the third limitation of this research, whose aim was to explicate and implement CPT: every analysis was programmatic in nature (as pointed out in Marlberg 2015: 354). It showed that a certain phenomenon existed in the language, and how it could be discovered. Still, research in this form is a compilation of fragments rather than an in-depth exploration of one particular research question.

The fragmentary nature of this research is mainly due to practical considerations. Textual segments were chosen that seemed to lend themselves to analysis more easily than others. Poems needed to be explored in full in order to prove that semantic auras interact

within a text, but that has left no room for deliberate comparison of particular linguistic aspects within or between languages. For example, metaphors are explored only in Yeats and only in one of his poems. Personification is only explored in Russian and the analysis comes from one text. The forensic aspect is studied only in Russian and is focused on newspaper discourse, because a forensic report was already available to the author. Prospection in novel writing is studied only in English, and only in American modernist and post-modernist prose. Pushkin takes the blame for insincerity at the level of subtext, but a parallel English example has not been found.

Still, it may be said that the research presents a variety of contexts in two languages, including poetry, novels, students' essays and newspaper discourse, revealing and highlighting consistency of findings when it comes to deploying the mechanisms outlined in Louw's publications. The consistency, as has been pointed out, is such that it may be said to verge on the unbelievable. As previously mentioned, the existence of relexicalisation and corpus-derived subtext has been illustrated in both languages. Counter-examples have not been found, and there was only one example reported of inconclusive findings (in Section 13.5 there was one example of a grammar string whose subtext led in various directions). This is all the more interesting since one of the texts studied was definitely not high-quality - the Russian manipulative newspaper excerpt, also containing a statement given by the victim of false rumours, who was definitely not a woman of letters. Still, both her excerpt and the newspaperman's were found consistent in their semantic auras. This consistency in various types of discourse provides strong evidence of CPT's reliability as a set of procedures aimed at extracting meaning from texts.

Using co-selection and wildcarding in order to explore semantic auras of an author's lexico-grammatical collocations in a balanced and representative reference corpus of the author's language must be accompanied by taking into account his or her usages in a reliable corpus of the author's texts. The researcher must also consider the context of situation and culture when studying each collocation, be it authorial or yielded by the reference corpus.

The remainder of this section will be devoted to the discussion of the research questions and objectives of this thesis as stated in Section 1.3.

'Is it possible to use reference corpora as a language norm against which particular authorial texts can be interpreted? If the answer to this question is positive, then is it legitimate to resort to inductive reasoning in textual interpretation?' The answer to these questions is positive for the following reasons. Reference corpora are balanced and representative and therefore may be viewed as a language norm, which is a basic tenet of the Sinclairan school and a modern principle of creating dictionaries. Particular authors are members of the language community and therefore deviations in their writing must be 'measured' against the same language norms that govern linguistic expression in this community. As for interpretation, every member of the language community conveys and receives meaning in accordance with the language norm that exist in their mind. The reference corpus is merely an objectivisation of it, containing language patterns representative of those that an adult native speaker of the language must have encountered in their lifetime. The consistency of semantic auras within texts as shown in Section 4.5 and subsequent sections is such that makes negating this impossible. Arguably, a particularly powerful proof of this could be prospection in novel writing – future research ought to focus on supplying more examples of this phenomenon. In philosophical terms, explained in Section 15, grammar is logical form and its most frequent lexis makes it 'perfect' (complete) language, while less frequent lexis used by the author deepens the meaning of an utterance. Inductive reasoning has therefore been proven as a logical step in inferring textual meaning from reference corpora, with the reservation that corpora must become much larger in the future without losing their representativeness.

'In connection with this, can semantic prosody be explained and defended as a mechanism underlying the predictive quality of language?' Semantic prosody has been shown to operate in the minds of human subjects in the psychological studies referred to in Section 4.4. However, it is not necessary to resort to psycholinguistic experimentation. If by 'predictive quality' we mean the ability of semantic prosody or corpus-derived subtext to indicate the likely outcome of a collocation *in a given context*, the answer appears to be affirmative. Flouted expectations result in nuances of meaning, as shown by the cline of prosodic and subtextual clashes in Section 9. Research has shown that deviations from the norm are not frequently so complete as to constitute a prosodic clash, but that discourse

comprises various kinds of deviation from the norm which involve both lexical and lexico-grammatical collocations and which should be explored in context. Poetic texts studied in sections 5, 7 and 15 show that celebrated poets employ collocations that are deeply indicative of semantic nuances. Statements by a lay person and a biased journalist (Section 8) also show remarkable consistency of semantic tendencies. However, if we follow Hunston (2011: 66) and say that by the 'predictive quality' it is meant that a prosodic clash must be understood as either expression of irony or insincerity, then research appears to confirm this statement. Lexico-grammatical collocations employed in the poem by Brodsky (Section 15.4) and the novel *Changing Places* by Lodge (Section 9.4) indicate that the persona and the character are being insincere - the ending to the poem by Brodsky proves this. These two cases seem to comprise authorial (i.e. typical of authors) insincerity and at the same time authorial irony (the author knows what the character does not). In the case of Pushkin (Section 5.4), authorial insincerity is not ironic. These examples confirm that every particular case should be thoroughly researched in its contextual setting. If semantic prosody is well-established by comprehensive research and an authorial collocation is in complete breach of this prevalent tendency, there must be some reason behind it, if not irony or insincerity: for example, Section 5.2.1 has shown that a diachronic change may also account for a prosodic mismatch. All in all, this research has not highlighted frequent cases of complete breaches of semantic prosody, in the sense that a prevalent tendency is completely reversed, but language has been shown to be highly predictive as to its meanings as revealed by the reference corpus, provided that collocations are viewed in their contexts of situation.

'How can the extraction of corpus-derived subtext contribute to nuanced textual interpretation?' Corpus-derived subtext is comprised of the most frequent lexical collocates of a grammar string and adds to the meaning of authorial collocation. It is completely opaque to intuition, as shown in Section 5, but can be part of rhetorical devices, as shown in Section 7. Its quality of pointing to further developments in texts, as shown in sections 4.5, 13 and 15, provides evidence that it is an objective phenomenon. Moreover, it can reveal authorial intention in conveying meaning (Section 5), and help to establish the

distinction between authorial, non-manipulative and manipulative insincerity (sections 5, 8 and 9).

'Does deviation from the reference corpus norm result either in a rhetorical device or in insincerity?' As stated above, deviation from the corpus norm takes many forms. It is illustrated by the cline of prosodic and subtextual clashes in Section 9, which shows that meaning is usually a combination of intentional and inadvertent aspects. Rhetorical devices and insincerity in its crudest form have been shown to occupy diametrically opposite points on the cline, which in fact has been shown to turn into a circle through the use of cliché either as a form of banality of manipulation. However, if by rhetorical device we mean a deviation that produces intended meaning, and by insincerity what is termed in sections 5 and 9 'authorial insincerity' or any other kind of forced expression or (non-)manipulative form of insincerity, then the answer to this question is positive.

'How can corpus-attested definitions of rhetorical devices be obtained?' The machanism of arriving at such definitions is shown in Section 7. Normally the questions asked would refer to the presence or absence of relexicalisation/delexicalisation, of a provisional meaning arrived at through co-selection, and of any indicative corpus-derived subtext. However, such research is at its earliest beginnings.

'What is the precise connection between Russel's and Wittgenstein's logical language and natural language use as studied by Contextual Prosodic Theory? How can this link be illustrated?' This link is illustrated in Section 15, which claims that in natural language vocabulary determines the speaker's attitude to the truthfulness of their own proposition (manipulative discourse excluded). The section claims that Russell refers to this phenomenon as 'facts asserted or denied', and Wittgestein as 'language games', in which 'there is an alteration – a gradual one – in the use of the vocabulary of a language' (OC 63).

'This dissertation will aim to present Contextual Prosodic Theory as a rounded and logically coherent theory, with an independent set of tools (methods and terminology), capable of textual interpretation without the help of another theory, e.g. from the field of pragmatics.' The claim that is defensible by this thesis is that CPT is in possession of a full descriptive mechanism which is necessary to formally and objectively interpret the meaning of a textual segment. The tools are, as stated, co-selection and wildcarding, and

the descriptive vocabulary includes 'collocation', 'relexicalisation', 'delexicalisation', 'semantic prosody', 'corpus-derived subtext', 'quasi-propositional variables', 'context of situation', 'context of culture', all of which have been given their definitions at appropriate junctures. Contextual Prosodic Theory is a corpus-stylistic theory that does not make use of terminology belonging to other fields of expertise, such as pragmatics or cognitive linguistics, but through its ability to arrive at textual interpretation it may verify the findings of these disciplines, which is a requirement included by Kozhin (Кожин 2010: 210) in his description of a rounded scientific system. His other requirement - that the introduction of new knowledge into a scientific system may not flout its inherent logic, but may only serve to sharpen existing terminology and introduce new terms into the already existing system – has also been fulfilled. New terminology (e.g. 'authorial intention', 'authorial insincerity') has been introduced without flouting the main logical system or application of CPT.

The implications of the findings if this dissertation will be discussed in the Conclusion.

17 Conclusion

As explained in Section 1.2 and Section 2, Contextual Prosodic Theory is a corpus stylistic theory which was first formally introduced to the academic community by Louw (2000), and then developed by him further in his subsequent publications to date. This thesis may be considered to have explained this theory in those areas which have been left insufficiently elaborated by its originator, and to have offered numerous practical examples of its application, given that Louw's contributions have largely been written from a theoretical perspective. This thesis has aimed to show that Contextual Prosodic Theory is an independent theory with a set of tools and procedures at its disposal, capable of extracting meaning from texts through consulting large and representative reference corpora. Although laborious and requiring special training, this kind of analysis does not require any special equipment or elaborate software, except for a computer and a reliable reference corpus. Should a specialised corpus become a necessity, for example of a particular author, such a corpus can easily be built. Nevertheless, as stated in the previous sections, the findings obtained by CPT methods are often not only opaque to intuition, but also unusually surprising. These facts encourage further speculation as to the scope of theory's implications and applicability.

Generally speaking, the findings may be said to lead to an impasse. On the one hand, the elaborateness of the interaction of semantic auras within texts (see Section 7 for examples from a poem by W. B. Yeats, and sections 7.5 and 15.3 from a poem by Brodsky) hints at a higher degree of the organisation of language than could previously have been assumed, even if one takes into account the outstanding talent of the authors in question. Talent as such, however, does not seem to be a prerequisite: the speaker and the writer whose texts were studied in Section 8 (a victim of gossip and a journalist writing disparaging texts on local politicians in a minor newspaper) cannot be said to possess any such talent, and yet the semantic auras in their usages exhibit clear consistencies. This higher degree of organisation in itself is in need of further research and could provide

valuable insight into cognitive mechanisms that are studied by disciplines other than linguistics, and in theory could lead to establishing disciplines that do not yet exist.

On the other hand, encouraging results of this sort may also be said to highlight an inconsistency: such a high organisation in discourse unfolds with less conscious input from its human creators than could previously have been imagined. It could even be stated that a higher organisation than was thought possible is continuously constructed by mechanisms that are more subliminal than could be suspected. Very tentatively, and perhaps wrongly, one could conclude that more meaning construction and construal might be taking place on the subliminal level than on the conscious level, including works of art that demand much conscious effort. For example, Brodsky builds a consistent denial of his own persona's apprehension of death (Section 15.3) in the first half of his poem, and then a continuous gradation of the degree of personification of pain. These coherent consistencies he could not have employed consciously, merely because the poem was written prior to the advent of computers, and yet they form a narrative device. This impasse could only be solved by an understanding that the human brain possesses capacities that are still to be investigated.

On a more practical level, much is to be gained by studying the layers of textual meaning that are not accessible to human intuition. In the field of literary criticism, for example, judgment as to the quality of a text may be arrived at more empirically. This is not to say that artistic effects are to be subjected to mechanical comparisons with the language norm - rather, they can now be more reliably explained. Moreover, texts may now be produced with assistance derived from reference corpora.[53] Other literary and translation-related applications have been summarised in the previous section.

Subliminal layers of meaning may also be of use in psychoanalysis. As hinted at in Section 15.5, if the language of poets and novel writers gives away how the poem/novel will end, any individual's discourse may be studied as to its consistent semantic auras, tell-tale signs of their general attitude to life's challenges and concerns. This 'predictive' power of language may then be used to confront the patient with their 'deviations from the norm',

[53] I am grateful for feedback I received from the audience at the Institute of Philosophy and Social Theory in Belgrade, to the effect that poems can now be produced using reference corpora, with prosodic clashes consciously exploited for the purposes of creating additional artistic effects.

the psychological 'norm' establishable through an appropriate reference corpus. If poets such as Pushkin and Larkin (Milojkovic 2011a) have been known to express a struggling and yet unrealised insight by verbal means, this means of analysis is open to every individual. This might call for the establishment of a 'forensic psycholinguistics'.

Insincerity in all its forms is a huge potential area of study. Manipulative discourse may be dismantled - and created - using CPT methodology. Louw (personal communication) is hopeful, however, that dismantling will always be more successful, as it is quite impossible to create a grammar string that will lie and not be highlighted by the computer. Still, since the general public does not normally resort to computers, the danger of manipulators using techniques such as those of CPT is not to be ignored. Forensically it is significant that, should findings of Section 8 be confirmed on a wider variety of examples, it will be possible to distinguish between intentional manipulators and those who primarily wish to deceive themselves. Not only forensic, but also psychological benefits of this are obvious.

The element of surprise that seems to accompany findings of CPT must also have a bearing on the approach to corpus investigations. Any 'rigorously formed hypothesis' (see Section 1.1) must allow for the possibility of its later modification, given that what the reference corpus may yield is often highly unpredictable.

A reservation must be expressed at this point. Section 12 has shown that even gifted and motivated students of philology, eager to arrive at strictly scientific findings, may commit grave errors of judgment when applying CPT methodology. Instruction in this area must be thorough and long-term, possibly preceded by a strict selection of candidates.

Generally speaking, since language pervades all areas of study, CPT methodology may be applied to a variety of issues. All such applications may now be impossible to predict. This points to the need to educate future researchers and the general public. Digital literacy (personal communication by Louw) must be the next step. Ironically, everyday experience shows that with the advent of computers analogue literacy has suffered. The findings of CPT point to the need to ensure that children keep being acquainted with text in all its forms, including the input by talented authors. The findings of this thesis show that it is the gifted author, poet or novel writer, that carries and perpetuates a language's meanings.

Analogue literacy must not be endangered by the advent of new technologies, but should be strengthened by a continuous computer-assisted study of its underlying mechanisms.

18 References

Abrams, M. H. and Harpham G. G. 2009. *A Glossary of Literary Terms*. Holt, Rinehart and Winston: New York

Albright, D. 1991. *W. B. Yeats: The Poems.* New York: Simon and Schuster Inc.

Alsuweed, M. 2015. *Accessing Dickens's Style as an EFL Learner: A Corpus Stylistic Approach to Lexical Style*. Unpublished doctoral thesis, University of Huddersfield (available at http://eprints.hud.ac.uk/id/eprint/31480/).

Ajmer K. (ed.). 2009. *Corpora and Language Teaching*. Amsterdam: John Benjamins.

Ayer, A.J. 1956. *The Problem of Knowledge*. Harmondsworth: Pelican Books.

Baker P. 1997. Consistency and accuracy in correcting automatically tagged corpora. In *Corpus Annotation*, R. Garside, G. Leech and T. McEnery (eds), 245-250. Harlow: Longman.

Baker P. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.

Baker, P., Hardie A. and McEnery T. 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bednarek, M. 2008. Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*, 4(2): 119-139.

Bennett, A., Cook, P., Miller, J. and Moore, D. 1963 [1987]. Aftermyth of war. In *The Complete Beyond the Fringe*, Alan Bennett, Peter Cook, Dudley Moore and Jonathan Miller (eds), 72-78. London: Methuen.

Berber-Sardinha, T. 2000. Semantic prosodies in English and Portuguese. A contrastive study. *Cuadernos de Filología Inglesa* 9(1): 93–110.

Burke, M. 2004. Cognitive stylistics in the classroom. *Style* 38(4): 491–510.

Cantor, A.P. 1991. Adolf, we hardly knew you. In *New Essays on White Noise*, F. Lentricchia (ed.), 39-62. Cambridge: CUP.

Carter, R. 2004. Introduction. In J. Sinclair, *Trust the Text, Language, Corpus and Discourse* (1-6). London: Routledge.

Castiglione D. 2019. *Difficulty in Poetry: A Stylistic Model*. Palgrave Macmillan.

Chapman, S. and Clark B. (eds). 2014. *Pragmatic Literary Stylistics*. Houndmills: Palgrave.

Clancy, B. and O'Keeffe, A. 2015 Pragmatics. In *Cambridge Handbook of English Corpus Linguistics*, D. Biber and R. Reppen (eds), 235–251. Cambridge: Cambridge University Press.

Clark, B. 2014. Pragmatics and inference. In *The Cambridge Handbook of Stylistics*, P. Stockwell and S. Whiteley (eds), 300-314. Cambridge: Cambridge University Press.

Cowart D. 2008. DeLillo and the power of language. In *The Cambridge Companion to Don DeLillo,* J.N. Duvall (ed.), 151-165. Cambridge: Cambridge University Press.

Cuddon, J. A. and Habib M. A. R. 1999. *A Dictionary of Literary Terms and Literary Theory*. Oxford: Wiley-Blackwell.

Culpeper, J. 2009. Keyness: words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*, *International Journal of Corpus Linguistics* 14 (1): 29–59.

Dam-Jensen H. and Zethsen K. 2008. Translator awareness of semantic prosodies. *Target* 20 (2): 203-211.

Davies, Mark. (2004-) *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available online at http://corpus.byu.edu/bnc/

Davies, M. 2008-. The Corpus of Contemporary American English: 520 million words, 1990-present. Available online at http://corpus.byu.edu/coca/

Davies, M. 2010-. The Corpus of Historical American English (COHA): 400 million words, 1810-2009. Available online at http://corpus.byu.edu/coha/.

Davies, M. 2011-. Google Books Corpus. (Based on Google Books n-grams). Available online at http://googlebooks.byu.edu/

Dickens, C. 1993. *The Pickwick Papers*. Ware: Wordsworth Editions Limite.d

Douthwaite, J. 2000. *Towards a Linguistic Theory of Foregrounding*. Edizioni Dell'Orso: Alessandria.

Duvall, J.N. 2008. Introduction: The power of history and the persistence of mystery. In *The Cambridge Companion to Don DeLillo*, Duvall J.N. (ed.), 1-10. Cambridge: CUP.

Ebeling, S. O. 2013. Semantic prosody in a cross-linguistic perspective. In *Corpus Linguistics and Variation in English: Focus on Non-Native Englishes*, M. Huber

and J. Mukherjee (eds), vol. 13, online. Studies in Variation, Contacts and Change in English.

Eco, U. 2001. Irony as the defining principle of postmodernism. In *Postmodernism*, D. Maus (ed.), 43-48. San Diego: Greenhaven Press.

Fialho, O. 2007. Foregrounding and refamiliarization: understanding readers' response to literary texts. *Language and Literature*, 16(2): 105-123.

Firth, J. R. 1957. *Papers in linguistics1934-1951*. Oxford: Oxford University Press.

Fischer-Starcke, B. 2010. *Corpus Linguistics in Literary Analysis: Jane Austen and Her Contemporaries*. London: Continuum.

Fogal, G. G. 2015 Pedagogical stylistics in multiple foreign language and second language contexts: A synthesis of empirical research. *Language and Literature*, 24(1): 54-72.

Forest, R. 2007. *'Tell Us What You Really Think': Collocation And Discourse in an Intertextual Corpus.* Unpublished PhD thesis, Department of English and Communication, City University, Hong Kong.

Foucault, M. 1969 [1984]. What is an author? In *The Foulcault Reader*, P. Rabinow (ed.), 101-120. New York: Pantheon Books.

Frijda, N.H. 1986. *The Emotions*. Cambridge: CUP.

Fuqua, J. 2014. Semantic prosody: The phenomenon of 'prosody' in lexical patterning. *The Journal of Language Teaching and Learning* 2, 76-83.

Gabrielaros C., Torgerson E., Hoffmann, S. and Fox, S. 2010. A corpus-based socio-linguistic study of the indefinite article forms in London English. *Journal of English Linguistics* 38(1): 1-38.

Gellner, E. 1959. *Words and Things*. Harmondsworth: Penguin.

Gilquin, G. and Gries, St. Th. 2009. Corpora and experimental methods: a state-of-the-art review. *Corpus Linguistics and Linguistic Theory* 5(1): 1-26.

Glock, H.-J. 2005. *A Wittgenstein Dictionary*. Oxford: Blackwell.

Greenbaum, S. (ed.). 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon Press.

Grice, H. P. 1957. Meaning. In *The Philosophical Review* 66(3): 377-388. Duke University Press on behalf of Philosophical Review.

Grice, H. P. 1975. Logic and conversation. In *Syntax and Semantics, Volume III: Speech Acts*, P. Cole and J. Morgan (eds), 41-58. New York: Academic Press.

Gries, S. 2010. Corpus linguistics, theoretical linguistics, and cognitive/psycholinguistics: towards more and more fruitful exchanges. *ICAME* 15.

Hall, G. 2014. Pedagogical stylistics. In *The Routledge Handbook of Stylistics*, M. Burke (ed.), 239–252. London: Routledge.

Halliday, M. A. K. 1971. Linguistic function and literary style: an inquiry into the language of William Golding's *The Inheritors*. In *Literary Style: a Symposium*, S. Chatman (ed.), 330-65. New York: Oxford University Press.

Halliwell, S. 1986. *Aristotle's Poetics.* London: Duckworth.

Hardie, A. and McEnery, T. 2010. On two traditions in corpus linguistics, and what they have in common. *ICAME* 15: 384-394.

Hauser, D. J. and Schwarz, N. 2016. Semantic prosody and judgment. *Journal of Experimental Psychology General* 145(7): 882-896.

Hauser, D. J. and Schwartz, N. 2018. How seemingly innocuous words can bias judgment: Semantic prosody and impression formation. *Journal of Experimental Social Psychology*, *75*, 11-18.

Herman, D. 2008. Narrative theory and the intentional stance. *Partial Answers: Journal of Literature and the History of Ideas* 6(2): 233-260.

Hirsch, E. D. 1976. Objective interpretation. In *On Literary Intention*, D. N. de Molina (ed.), 26-54. Edinburgh University Press.

Hoey, M. 2005. *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hucheon, L. 2001. *The Politics of Postmodernism*. London: Routledge.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hunston, S. 2007. Semantic prosody revisited. *International Journal of Corpus Linguistics* 12 (2): 245-264.

Hunston, S. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London: Routledge.

Hunston, S. and Francis, G. 1999. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English.* Amsterdam: John Benjamins.

Jakobson, R. 1960. Closing statements: linguistics and poetics. In *Style in Language*, T. A. Sebeok (ed.), 350-377. Cambridge Massachusetts: MIT Press.

Jeffries, L. 2010. *Critical Stylistics*. Basingstoke: Palgrave.

Jeffries, L. and McIntyre, D. 2010. *Stylistics*. Cambridge: Cambridge University Press.

Jeffries, L., and Walker, B. 2017. *Keywords in the Press: The New Labour Years*. London: Bloomsbury.

Kershner, R. B. 1997. *The Twentieth-Century Novel: An Introduction*. Boston: Bedford Books.

Krishnamurthy, R. 2000. Collocation: from *silly ass* to lexical sets. In *Words in Context: A Tribute to John Sinclair on His Retirement*, C. Heffer and H. Sauntson (eds), 31-47. University of Birmingham.

Kübler, N. and Volanschi, A. 2012. Semantic prosody and specialised translation, or how a lexicogrammatical theory of language can help with specialised translation. In *Corpus-informed Research and Learning in ESP: Issues and Applications*, A. Boulton, S. Carter-Thomas, and E. Rowley-Jolivet (eds), 105-135. Amsterdam: John Benjamins.

Kučera, H. and Francis, W. N. 1967. *Computational Analysis of Present Day American English.* Providence, RI: Brown University Press.

Lakoff, G. and Johnson, M. 1980 [2003]. *Metaphors We Live by*. Chicago: The University of Chicago Press.

Laudan, L. 1981. *Science and Hypothesis.* Dordrecht: Springer.

Lettrucchia, F. 1991. Tales of the electronic tribe. In *New Essays on White Noise*, F. Lentrucchia (ed.), 87-113. Cambridge: CUP.

Lodge, D. 1975. *Changing Places*. London: Penguin Books.

Lodge, D. 1984. *Small World.* Harmondsworth: Penguin Books.

Lodge, D. 2000. *Thinks*. London: Secker and Warburg.

Louw, W. E. 1991. Classroom concordancing of delexical forms and the case for integrating language and literature. In *Classroom Concordancing*, T. Johns, and P.

King (eds), 151-178. Birmingham: Centre for English Language Studies, the University of Birmingham.

Louw W. E. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In *Text and Technology: In Honour of John Sinclair,* M. Baker, G. Fransis and E. Tognini-Bonelli (eds), 152-176. Amsterdam: John Benjamins.

Louw, W. E. 2000. Contextual prosodic theory: bringing semantic prosodies to life. In *Words in context*. *In honour of John Sinclair*, C. Heffer, and H. Sauntson (eds), 48-94. Birmingham: ELR.

Louw, W.E. 2003. Dressing up waiver: A stochastic collocational reading of 'The Truth and Reconciliation Commission (TRC). *Occasional Papers dei Quaderni del CeSLIC*. Bologna: University of Bologna. Available at https://uz-ac.academia.edu/BillLouw.

Louw, W. E. 2006. Data-assisted negotiating: Will it produce a new class of negotiator or destroy the ideology of negotiating? *Language Matters* 37(2).

Louw, W.E. 2007a. Collocation as the determinant of verbal art. In *Verbal Art Re-Visited*, D. Miller and M. Turci (eds), 149–181. London: Equinox.

Louw, W.E. 2007b. Truth, literary worlds and devices as collocation. In *Corpora in Foreign Language Classroom,* E. Hidalgo, L. Quereda and J. Santana (eds), 329–362. Amsterdam: Rodopi.

Louw, W.E. 2007c. Literary worlds as collocation. In *Literature and Stylistics for Language Learners: Theory and Practice*, G. Watson and S. Zyngier (eds), 91-105. Basingstoke: Palgrave.

Louw, W.E. 2008. Consolidating empirical method in data-assisted stylistics: Towards a corpus-attested glossary of literary terms. In *Directions in Empirical Literary Studies. In Honour of Willie van Peer* [Linguistic Approaches to Literature 5], S. Zyngier, M. Bortolussi, A. Chesnokova and J. Auracher (eds), 243–264. Amsterdam: John Benjamins.

Louw, W. E. 2009. The analysis and creation of humour. Keynote address to the Second International Conference on Humour Studies, Lucca, Italy.

Louw, W.E. 2010a. Collocation as instrumentation for meaning: A scientific fact. In *Literary Education and Digital Learning: Methods and Technologies for Humanities Studies*, W. van Peer, V. Viana and S. Zyngier (eds), 79–101. Hershey PA: IGI Global.

Louw, W.E. 2010b. Automating the extraction of literary worlds and their subtexts from the poetry of William Butler Yeats. In *Para por y sobre Luis Quereda,* M. Falces Sierra, E. Hidalgo Tenorio, J. Santana Lario and S. Valera Hernandez (eds), 635–657. Granada: Granada University Press.

Louw, W.E. 2011. Philosophical and literary concerns in corpus linguistics. In *Perspectives on Corpus Linguistics* [Studies in Corpus Linguistics 48], V. Viana, S. Zyngier and G. Barnbrook (eds), 171–196. Amsterdam: John Benjamins.

Louw, W. E. 2013. Teaching literary engagement from the index of first lines: a corpus-based subtextual approach. In *Proceedings of the International Conference 'New Techniques and Technologies in Teaching Foreign Languages, MGIMO, 4-5 October 2013.* Vol. 1. Крячков Д. А. (ed.), 340-344. Moscow: MGIMO University.

Louw, W. E. 2017a. Uneasy humour as discovery: Collocation and empathy as Whewellian consilience. *Studying Humour: International Journal*, Vol. 4.

Louw, W. E. 2017b. Entailing proximity. In *Труды международной конференции «Корпусная лингвистика – 2017»*, Zakharov V.P. (et al.), 53-58. Санкт-Петербург: С.-Петербургский гос. университет, Фил. факультет.

Louw, B. 2017c. Negotiators' language as code and cipher. In *Магия ИННО: новые технологии в языковой подготовке специалистов-международников*. (Москва, 25–26 марта 2017 г). Том 1, Секции 1-4. 89-96. Moscow: Moscow State University of International Relations MGIMO.

Louw, W.E. (in press). Towards a Theory of Corpus Linguistics: Proofs Banish Proscription. *Структурная и прикладная лингвистика* [межвуз. сб. вып. 13], отв.ред. И.С. Николаев. - СПб., Издательство С.-Петербургского университета.

Louw, W. E., and Milojkovic, M. 2014. Semantic Prosody. In *The Cambridge handbook of stylistics*, P. Stockwell, and S. Whiteley (eds), 263-280. Cambridge: Cambridge University Press.

Louw, W. E and Milojkovic, M. 2015. Shared logical form or shared metaphysics: In search of corpus-derived empathy in stylistics. In *7th International Conference on Corpus Linguistics: Current Work in Corpus Linguistics: Working with Traditionally-conceived Corpora and Beyond (CILC 2015)* [Procedia - Social and Behavioral Sciences 198], 535 – 545. Elsevier.

Louw, B. and Milojkovic, M. 2016. *Corpus stylistics as Contextual Prosodic Theory and subtext. Amsterdam:* John Benjamins.

Louw, W.E. and Milojkovic, M. 2016b.War Poets as a Brand – Unpacked by Corpus-Derived Subtext. In *CILC2016. 8th International Conference on Corpus Linguistics*, [EPiC Series in Language and Linguistics Volume 1], A. Moreno Ortiz and C. Pérez-Hernández (eds), 260-273.

Lugea, J. 2017. This year's work in stylistics 2016. In *Language & Literature* 26(4): 352-353.

Mahlberg, M. 2013. *Corpus Stylistics and Dickens's Fiction.* London: Routledge.

Mahlberg, M. 2015. Literary style and literary texts. In *Cambridge Handbook of English Corpus Linguistics*, D. Biber and R. Reppen (eds.), 235–251. Cambridge: Cambridge University Press

Malinowski, B. 1923. The problem of meaning in primitive languages. In *The Meaning of Meaning*, C. K. Ogden and I. A. Richards (eds), 296–336. London: Routledge.

Masevich, A. C. and Zakharov, V.P. 2016. Metody korpusnoj lingvistiki v istoriheskih i kul'torologicheskih issledovaniyah [Corpus Linguistics Methods in Historical and Cultural Studies] (https://www.academia.edu/29209174, accessed 19 November 2016)

Maus, D. 2001. (ed.). *Postmodernism*. San Diego: Greenhaven Press.

Mautner, T. 2000 [2005]. *The Penguin Dictionary of Philosophy*. London: Penguin.

McEnery, T. and Hardie, A. 2012. *Corpus Linguistics*. Cambridge: CUP.

McGee, I. 2012. Should we teach semantic prosody awareness? *RELC Journal* 43(2): 169-186.

McIntyre, D. 2018. Irony and semantic prosody revisited. In *The Pragmatics of Irony and Banter*, M. Jobert, and S. Sorlin (eds), 81-99. Amsterdam: John Benjamins.

Miall, D. S. and Kuiken, D. 1994. Foregrounding, defamiliarization and affect: response to literary stories. *Poetics* 22: 389 - 407.

Miall, D. S. and Kuiken, D. 2001. Shifting perspectives: readers' feelings and literary response. In *New Perspectives on Narrative Perspective,* W. van Peer, and S. Chatman (eds.), 289-301. Albany, NY: State University of New York Press.

Miall, D. S., and Kuiken, D. 2002. A feeling for fiction: becoming what we behold. *Poetics* 30: 221 - 41.

Miller, H. 1967. *Tropic of Capricorn.* London: Panther Books.

Milojkovic, M. 2011a. Quenched light, or seeing through a glass darkly – a collocation-based view of Larkin's atheism and depression. *Belgrade English Language and Literature Studies* 3: 127-144.

Milojkovic, M. 2011b. Semantic prosody and subtext as universal, collocation-based instrumentation for meaning and literary worlds. In *Труды международной конференции «Корпусная лингвистика – 2011»*, Zakharov V.P. et al. (ed.), 47-52. St Petersburg: St Petersburg State University, Faculty of Philology.

Milojkovic, M. 2012. Time and transitions in Larkin's poetry. In *NAWA: Journal of Language and Communication* 6(1): 102–126.

Milojkovic, M. 2013a. Is corpus stylistics bent on self-improvement? The role of reference corpora 20 years after the advent of semantic prosody. In *Journal of Literary Semantics* 42(1): 59–78.

Milojkovic, M. 2013b. Bill Louw's contextual prosodic theory as the basis of (foreign language) classroom corpus stylistics research. *Research in Corpus Linguistics* 1: 47-63.

Milojkovic M. 2016a. A corpus-attested view of business English metaphors. In *CILC2016. 8th International Conference on Corpus Linguistics*, [EPiC Series in Language and Linguistics, volume 1], A. Moreno Ortiz and C. Pérez-Hernández (eds), 283-297.

Milojkovic M. 2016b. Corpus-derived subtext in Russian-English translation. In *Прикладная лингвистика в науке и образовании. Alpac Report – полвека после разгрома. Труды VIII Международной научной конференции. 24-26 ноября*

2016 г., 188-193. Санкт-Петербург: Филологический факультет РГПУ им. А.И. Герцена.

Milojkovic M. 2017a. Grammatical 'hidden' meaning as part of pun and metaphor in English. *Studying Humour – International Journal* 4.

Milojkovic M. 2017b. The reconstruction of the context of culture through corpus stylistics. In *Труды международной конференции «Корпусная лингвистика – 2017»*, V.P. Zakharov (ed.), 59-65. Санкт-Петербург: С.-Петербургский гос. университет, Фил. факультет.

Milojkovic, M. (in press). Corpus-derived subtext and prospection in novel-writing: examining Faulkner's *Absalom, Absalom!* and DeLillo's *White Noise.* In *Структурная и прикладная лингвистика* [межвуз. сб. вып. 13], отв.ред. И.С. Николаев. - СПб., Издательство С.-Петербургского университета.

Milojkovic, M. and Louw, W.E. 2017. Towards a corpus-attested definition of creativity as accessed through a subtextual analysis of student writing. In *Key Competences in English Medium Higher Education*, R. Breeze and C. Sancho Guinda (eds), 125-141. Springer.

Mitchell, K. 2008. *Intention and Text: Towards an Intentionality of Literary Form*. London: Continuum.

Mitscherling J., DiTommaso T., and Nayed A. 2004. *The Author's Intention*. Lanham: Lexington Books.

Moon, R. 1998. *Fixed Expressions and Idioms in English: A Corpus-based Approach*. Oxford: Oxford University Press.

Moon, R. 2007. Words, frequencies and texts (particularly Conrad). *Journal of Literary Semantics* 36: 1–33.

Morley, J. and Partington A. 2009. A few frequently asked questions about semantic - or evaluative - prosody. *IJCL* 14:2. 140-158.

Morrison, T. 1994. *Conversations with Tony Morrison*, D. Taylor-Guthrie (ed.). Jackson: University Press of Mississippi.

Mukařovský, J.1970. Standard language and poetic language. In *Linguistics and Literary Style,* D. C. Freeman (ed.). New York, NY: Holt, Rinehart and Winston.

Mukherjee, J. 2005. Stylistics. In *Encyclopaedia of Linguistics*, P. Strazny (ed.), 1184–6. New York: Fitzroy Dearborn.

O'Halloran, K. 2007a. Corpus-assisted literary evaluation. *Corpora* 2(1), 33-63.

O'Halloran, K. 2007b. Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics* 28: 1–24.

Olster, S. 2008. White Noise. In *The Cambridge companion to Don DeLillo*, J.N. Duvall (ed.), 79-93. Cambridge: CUP.

Oster, U. 2010. Using corpus methodology for semantic and pragmatic analyses: what can corpora tell us about the linguistic expression of emotions? *Cognitive Linguistics* 21(4): 727–763.

Partington, A. 2004. 'Utterly content in each other's company': semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1): 131–156.

Partington, A. 2014. Evaluative prosody. In *Corpus Pragmatics: a Handbook*, K. Ajmer and C. Rüchlemann (eds.), 279-303. Cambridge: Cambridge University Press..

Pears, D. 1971. *Wittgenstein.* London: Fontana Collins.

Philip, J. 2011. *Colouring Meaning. Collocation and Connotation in Figurative Language*. Amsterdam: John Benjamins.

Pushkin, A. 2009. *Selected Lyric Poetry*. Trans. J.E. Falen. Evanston, Illinois: Northwestern University Press.

Quirk, R., Greenbaum S., Leech G. and Svartvik J. 1972. *A Grammar of Contemporary English.* London: Longman.

Quirk, R., Greenbaum S., Leech G. and Svartvik J. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Reichenbach, H. 1944 [1998]. *The Philosophy of Space and Time*. New York: Dover Books.

Ruppersburg, H. and Engles T. 2000. Introduction. In *Critical Essays on Don DeLillo*, H. Ruppersburg and T. Engles (eds), 1-28. New York: G. K. Hall and Co.

Russell, B. 1946. *The History of Western Philosophy.* London: Routledge.

Russell, B. 1956 [2007]. *My Philosophical Development.* Nottingham: Spokesman Press.

Russell, B. 1960. *Bertrand Russell speaks his mind*. London: Arthur Barker.

Russell, B. 1996. *An Inquiry into Meaning and Truth*. Routledge: London

Sarker, S. K. 2002. *W. B. Yeats: Poetry and Plays.* New York: Atlantic Publishers and Distributors.

Semino, E. and Short, M. 2004. *Corpus Stylistics. Speech, Writing and Thought Presentation in a Corpus of English Writing.* London: Routledge.

Shklovsky, V. 1919. Art as device. Downloaded from https://warwick.ac.uk/fac/arts/english/currentstudents/undergraduate/modules/fulllist/first/en122/lecturelist2017-18/art_as_device_2015.pdf (accessed in March 2019).

Siepmann, D. 2015. A corpus-based investigation into key words and key patterns in post-war fiction. *Functions of Language* 22(3): 362–399.

Simpson, P. 2014. *Stylistics*. London: Routledge.

Sinclair, et al. (eds). 1990. *Collins COBUILD English Grammar*. London: HarperCollins Publishers.

Sinclair, J. M. 1991. *Corpus, Concordance, Collocation.* Oxford: OUP.

Sinclair, J. M. 1996. The search for units of meaning. *Textus* IX. 75-106.

Sinclair, J. M. 2003. *Reading Concordances.* London: Longman.

Sinclair, J. M. (edited with Ronald Carter). 2004. *Trust the Text. Language, Corpus and Discourse*. London: Routledge.

Sinclair, J.M. 2006. *Phrasebite*. Pescia: TWC

Sinclair, John McH. 2010. Defining the definiendum. In *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks,* G.M. de Schryver (ed.), 37 – 47. Uganda, Kampala: Menha Publishers Ltd.

Sinclair, J. M. and Renouf A. J. 1987. A lexical syllabus for language learning. In *Vocabulary and Language Teaching,* M. J. McCarthy and R. A. Carter (eds), 140-160. London: Longman.

Sinclair, J. M., Jones S. and Dalwy R. 2004. *English Collocation Studies: The OSTI Report,* R. Krishnamurthy (ed.). London: Continuum.

Sotirova, V. 2015. Production and intentionality. In *The Cambridge Handbook of Stylistics*, P. Stockwell and S. Whiteley (eds.), 132-148. Cambridge: Cambridge University Press.

Stanford Encyclopedia of Philosophy. Retrieved from https://plato.stanford.edu/ (accessed in July 2017).

Steen, G. 2015. Metaphor and style. In *The Cambridge Handbook of Stylistics*, P. Stockwell and S. Whiteley (eds), 315-328. Cambridge: Cambridge University Press.

Stefanowitsch, A. (forthcoming). *Corpus Linguistics: A Guide to the Method*. Language Science Press.

Stewart, D. 2009. Safeguarding the lexicogrammatical environment: Translating semantic prosody. In *Corpus Use and Translating: Corpus Use for Learning to Translate and Learning Corpus Use to Translate*, A. Beeby, P. Rodriguez-Ines and P. Sanchez-Gijon (eds.), 29-46. Amsterdam: John Benjamins.

Stewart, D. 2010. *Semantic Prosody. A Critical Evaluation*. New York/London: Routledge.

Stockwell, P and Mahlberg, M. 2015. Mind-modelling with corpus stylistics in *David Copperfield*. *Language and Literature* 24, 129–147.

Stockwell, P. 2002. Cognitive Poetics. London: Routledge.

Stockwell, P. 2009. *Texture: A Cognitive Aesthetics of Reading.* Edinburgh: Edinburgh University Press.

Stockwell, P. 2015. Poetics. In *Handbook of Cognitive Linguistics* [Handbooks of linguistics and communication science (39)], E. Dabrowska and D. Divjak (eds), 432-452. de Gruyter Mouton.

Stubbs, M. 1995. Collocations and semantic profiles: on the cause of the trouble with quantitative methods. *Function of Language* 2(1): 1–33.

Stubbs, M. 2001. *Words and Phrases*. Oxford: Blackwell.

Stubbs, M. 2005. Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature* 14:1, 5–24.

Stubbs, M. 2014. Quantitative methods in literary linguistics. In *The Cambridge Handbook of Stylistics*, P. Stockwell and S. Whiteley (eds), 46–62. Cambridge: Cambridge University Press.

Svartvik, J. (ed.). 1990. *The London-Lund Corpus of Spoken English: Description and Research*. Malabar, FL: Krieger Publishing Company.

Tabata, T. 2002. Investigating stylistic variation in Dickens through correspondence analysis of word-class distribution. In *English Corpus Linguistics in Japan*, T. Saito, J. Nakamura and S. Yamazaki (eds.), 165–182. Amsterdam, New York: Rodopi.

Teubert, W. 2005. My version of corpus linguistics. *International Journal of Corpus Linguistics* 10(1): 1-13.

Teubert, W. 2010. *Meaning, Discourse and Society*. Cambridge: Cambridge University Press.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.

Toolan, M. 2007. Trust the text, text as trust. *IJCL* 12:2, 269-288.

Toolan, M. 2009. *Narrative Progression in the Short Story: A Corpus Stylistic Approach.* Amsterdam: John Benjamins.

Toolan, M. 2018. Stylistics. In *A Companion to Literary Theory* [Blackwell Companions to Literature and Culture], D. H. Richter (ed.), 61-71. Oxford: Wiley-Blackwell.

Village, A. 2007. *The Bible and Lay People: an Empirical Approach to Ordinary Hermeneutics*. Aldershot: Ashgate.

Wang, F.R. and Humblé P. 2017. Corpus Stylistics as Contextual Prosodic Theory and Subtext by Bill Louw, Marija Milojkovic. *Style* 51(4): 550-555.

Wei, N. and Li, X. 2014. Exploring semantic preference and semantic prosody across English and Chinese: their roles for cross-linguistic equivalence. *Corpus Linguistics and Lingustic Theory* 10(1): 103-138.

Whitsitt, S. 2005. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics* 10: 283-305.

Williams, G. 2010. Many rooms with corpora. *ICAME* 15, 400-407.

Wimsatt, W. K. and Beardsley M. C. 1954. The intentional fallacy. In *The Verbal Icon: Studies in the Meaning of Poetry*, W. K. Wimsatt (ed.), 2-18. Lexington: University of Kentucky Press.

Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*. Trans. D.F. Pears and D.F. McGuiness, 1960. London: Routledge and Kegan Paul.

Wittgenstein, L. 1929. Some remarks on logical form. *Proceedings of the Aristotelian Society, Supplementary Volumes,* Vol. 9, Knowledge,Experience and Realism, 162-171. Blackwell Publishing on behalf of The Aristotelian Society http://www.jstor.org/stable/4106481 (accessed in April 2019).

Wittgenstein, L. 1969. *On Certainty*, G.E.M. Anscombe and G.H. von Wright (eds). New York: Harper and Row. Retrieved from http://thatmarcusfamily.org/philosophy/Course_Websites/Readings/Wittgenstein%20On%20Certainty.pdf (accessed in July 2017).

Xiao, R. Z. 2013. Review of McEnery and Hardie (2012) Corpus Linguistics: Method, Theory and Practice. *Corpus Linguistics and Linguistic Theory* 9(2): 293-303.

Xiao, R. Z. and McEnery, T. 2006. Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics* 27: 103-129.

Yeats, W. B. 1939. *Last Poems*. London: Macmillan.

Zyngier, S. and Fialho, O. 2010. Pedagogical stylistics, literary awareness and empowerment: a critical perspective. *Language and Literature* 19(1): 13–33.

Иссерс, О.С. 2008. *Коммуникативные стратегии и тактики русской речи*. М.: Едиториал УРСС.

Кожин, П.М. 2007. *Этнокультурные контакты населения Евразии в энеолите — раннем железном веке (палеокультурология и колесный транспорт*). Владивосток: Дальнаука.

Милойкович, М. (in press) 2019a. Возможности контектуально-просодической теории для установления диагностических признаков манипулятивного дискурса при производстве лингвистической экспертизы. Acta linguistica Petropolitana ― *Труды Института лингвистических исследований РАН.*

Милойкович, М. (in press) 2019b. Иосиф Бродский: модальность как прием. *Вестник Нижегородского университета им. Н.И. Лобачевского*.

Радбиль, Т.Б. 2014. Выявление содержательных и речевых признаков недобросовестной информации в экспертной деятельности лингвиста. *Вестник Нижегородского университета им. Н.И. Лобачевского* 6: 146-149.

Радбиль, Т.Б., Юматов В.А. 2014. Способы выявления имплицитной информации в лингвистической экспертизе. *Вестник Нижегородского университета им. Н.И. Лобачевского* 3(2): 18-21.

19 Биографија аутора

Марија Милојковић је рођена 14.12.1975 у Руској Федерацији. Завршила је Средњу школу при Амбасади Руске Федерације у СР Југославији 1992. године и Филолошки факултет Универзитета у Београду, програм Енглески језик и књижевност, 1997. године. Радила је као професор енглеског језика у Школи sa стране језике Задужбине Илије М. Коларца између 1998. и 2000. године и као асистент за предмете Енглеска књижевност I и II на Катедри за англистику Филолошко-уметничког факултета Универзитета у Крагујевцу између 2000 и 2002. године. Запослила се на Катедри за англистику Филолошког факултета Универзитета у Београду 2002. године као лектор за Савремени енглески језик, где је унапређена у вишег лектора 2018. године.

Завршила је мастер студије 2010. године, уписала докторске студије први пут 2011, и други пут 2015. године. Бави се корпусном стилистиком од 2010. године. Наступала је на конференцијама из области корпусне лингвистике и стилистике у Енглеској, Италији, Намибији, Немачкој, Русији, Словенији, Шпанији, и позвана је да држи пленарни наступ на међународној конференцији из области судске експертизе у организацији Универзитета „ Лобачевски" у Нижњем Новгороду, РФ, која ће се одржати у мају 2019. године. Одржала је два предавања по позиву: на Универзитету у Модени, Италија, у новембру 2017. године, и на Институту за Филозофију и друштвену теорију у Београду у децембру 2018. године. Аутор је више научних чланака на енглеском и руском језику и књиге на енглеском језику у коауторству с Билом Лоуом, која је изашла 2016. године у издању куће Џон Бенџаминс у Амстердаму.

Члан је међународне асоцијације Poetics and Linguistics Association.

Објављени радови

Књига:

Louw, B. and Milojkovic, M. (2016). Corpus Stylistics as Contextual Prosodic Theory and Subtext. Amsterdam: John Benjamins

Поглавља у књигама:

Milojkovic M. (in press). Corpus-derived subtext and prospection in novel-writing: examining Faulkner's *Absalom, Absalom!* and DeLillo's *White Noise.* In *Структурная и прикладная лингвистика* [межвуз. сб. вып. 13], отв.ред. И.С. Николаев. - СПб., Издательство С.-Петербургского университета.

Milojkovic M. and Louw W. E. 2017. Towards a corpus-attested definition of creativity as accessed through a subtextual analysis of student writing. In *Essential Competencies for English-medium University Teaching*, R. Breeze and C. Sancho Guinda (eds), 125-141. Springer.

Louw W. E. and Milojkovic M. 2014. Semantic prosody. In *The Cambridge Handbook of Stylistic*, P. Stockwell and S. Whiteley (eds), 260-283. Cambridge: Cambridge University Press.

Чланци у часописима:

Милойкович М. (принято в печать) 2019a. Возможности контектуально-просодической теории для установления диагностических признаков манипулятивного дискурса при производстве лингвистической экспертизы. Acta linguistica Petropolitana ― Труды Института лингвистических исследований РАН.

Милойкович, М. (принято в печать) 2019b. Иосиф Бродский: модальность как прием. Вестник Нижегородского университета им. Н.И. Лобачевского.

Milojkovic M. 2017a. Grammatical 'hidden' meaning as part of pun and metaphor in English. *Studying Humour – International Journal* 4

Milojkovic M. 2013a. Is corpus stylistics bent on self-improvement? The role of reference corpora 20 years after the advent of semantic prosody. *Journal of Literary Semantics* 42 (1): 59–78.

Milojkovic M. 2013b. Bill Louw's Contextual Prosodic Theory as the basis of (foreign language) classroom corpus stylistics research. *Research in Corpus Linguistics* 1: 47-63.

Milojkovic M. 2012. Time and transitions in Larkin's poetry. *NAWA: Journal of Language and Communication* 6 (1): 102-126.

Milojković M. 2011a. Quenched light, or seeing through a glass darkly – a collocation-based view of Larkin's atheism and depression. *Belgrade English Language and Literature Studies* 3: 127-144.

Чланци у зборницима с конференција:

Milojkovic M. 2017b. The reconstruction of the context of culture through corpus stylistics. In *Труды международной конференции «Корпусная лингвистика – 2017»,* V.P. Zakharov (ed.), 59-65. Санкт-Петербург: С.-Петербургский гос. университет, Фил. факультет.

Милойкович М. 2017с. „Возможности описания культурных контекстов методом корпусной стилистики". In *Магия ИННО: новые технологии в языковой подготовке специалистов-международников*. *(Москва, 25–26 марта 2017 г)*. Том 2, Секции 5-8. 536-544.

Milojkovic M. 2016a. A corpus-attested view of business English metaphors. In *CILC2016. 8th International Conference on Corpus Linguistics* [EPiC Series in Language and Linguistics, volume 1], A. Moreno Ortiz and C. Pérez-Hernández (eds), 283-297..

Louw W.E. and Milojkovic M. 2016. War poets as a brand – unpacked by corpus-derived subtext. In *CILC2016. 8th International Conference on Corpus Linguistics* [EPiC Series in Language and Linguistics, volume 1], A. Moreno Ortiz and C. Pérez-Hernández (eds), 260-273.

Milojkovic M. 2016b. Corpus-derived subtext in Russian-English translation. In *Прикладная лингвистика в науке и образовании. Alpac Report – полвека после разгрома*. *Труды VIII Международной научной конференции. 24-26 ноября 2016 г.*, 188-193. Санкт-Петербург: Филологический факультет РГПУ им. А.И. Герцена.

Louw W. E and Milojkovic M. 2015. Shared logical form or shared metaphysics: In search of corpus-derived empathy in stylistics. In *7th International Conference on Corpus*

Linguistics: Current Work in Corpus Linguistics: Working with Traditionally-conceived Corpora and Beyond (CILC 2015) [Procedia - Social and Behavioral Sciences 198], 535 – 545. Elsevier.

Kalanj B. and Milojkovic M. 2014. Semantic prosody and subtext in the writing samples of Belgrade students of English. On-line Proceedings of the Annual Conference of PALA. http://www.pala.ac.uk/Annual conferences/proceedings/2014/kalanj-milojkovic2014.doc

Milojkovic M. 2013c. Semantic prosody and subtext in English-Russian poetic translation. In *Proceedings of the International Conference 'The Magic of Innovation: New Techniques and Technologies in Teaching Foreign Languages'*, 411-415. Moscow: MGIMO University.

Milojković M. 2011c. The role of teaching techniques in vocabulary acquisition. In *ELLSIIR Proceedings. Volume 1*, N. Tomović and Vujić J. (eds), 543-555. Belgrade: Faculty of Philology, University of Belgrade.

Milojkovic M. 2011b. Semantic prosody and subtext as universal, collocation-based instrumentation for meaning and literary worlds. In *Труды международной конференции «Корпусная лингвистика – 2011*, V.P. Zakharov (ed.), 47-52. Санкт-Петербург: С.- Петербургский гос. университет, Филологический факультет.

Преглед књиге:

Milojkovic M. 2015. Michaela Mahlberg, Corpus Stylistics and Dickens's Fiction. In *English Text Construction* 8 (1): 143–152.

Изјава о ауторству

Име и презиме аутора Марија Милојковић

Број индекса 15083/А

Изјављујем

да је докторска дисертација под насловом

Експликација и примена Лангове корпусно-стилистичке конекстуално-прозодијске теорије на материјалима енглеског и руског језика

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 14.05.2019.

Марија Милојковић

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора: Марија Милојковић

Број индекса: 15083Д

Студијски програм: Језик, књижевност, култура (модул: језик)

Наслов рада: Експликација и примена Лоуове корпусно-стилистичке контекстуално-прозодијске теорије на материјалима енглеског и руског језика

Ментор: проф. др Јелена Вујић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, 14.05.2019.

Марија Милојковић

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Експликација и примена Лоуове корпусно-стилистичке контекстуално-прозодијске теорије на материјалима енглеског и руског језика

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 14.05.2019.

Марија Милојковић