RiCL

Research in Corpus Linguistics

Bill Louw's Contextual Prosodic Theory as the basis of (foreign language) classroom corpus stylistics research

Marija Milojkovic[1]
University of Belgrade / Serbia

Abstract – Corpus empiricism may alter the act of reading. This began as the reader searched a reference corpus for individual words and phrases. With the admission of lexicographers that intuition no longer suffices in providing a definition, corpus stylistics must go further by showing that a literary text can no longer be properly interpreted if not seen against the background of the wealth of recorded textual experience. This by no means suggests that a literary text may not have a satisfying impact on an individual reader; rather, corpus stylistics enhances our interpretation by means that are easily available. The core of Bill Louw's stylistic approach is his claim that prior knowledge is no longer perceived as concepts (unsatisfyingly intuitive).

Therefore, reference corpora may serve to enhance our stylistic interpretation of a literary text that was clearly written to be appreciated as a unique textual experience. Roughly, a large reference corpus will provide many parallel textual experiences, so that 'events' in the studied text are augmented by their counterparts in corpora. Thus, our understanding of the text will be augmented by what is absent from it, but present in the reference corpora. If, furthermore, our classroom is a foreign language one, the reference corpus will serve as missing language experience in the foreign language learner, even if the learner is very proficient.

After giving a brief overview of Louw's Contextual Prosodic Theory (CPT) and its implications for classroom corpus stylistics, the paper describes a study conducted with second-year students of English from the University of Belgrade. The aims of the study are to verify Louw's principle that text reads text and to test the proposed CPT-based methodology. The study consists of a quantitative part (where the learning phase is followed by a final test) and a qualitative part (questionnaire). The proposed methodology relies on confronting the subjects with concordance lines as a means of interpreting a collocation in a given short excerpt, with an absolute minimum of theoretical background. The subjects are tested on semantic prosodies, absent collocates and auras of grammatical strings, through tasks that vary in format. The results obtained are encouraging for CPT, despite the study's limitations, which are also discussed.

Keywords – Bill Louw, events, methodology, Contexual Prosodic Theory, corpus stylistics, semantic prosody, subtext

1. INTRODUCTION

1.1. The aim and significance of the study

The aim of this paper is to present a piece of classroom stylistics research based on the theoretical implications of Contextual Prosodic Theory (CPT), developed by Bill Louw from 1993 to date. As this is the first instance of practical implementation of CPT for the purposes of teaching corpus stylistics, the study will attempt to answer two research questions that are interrelated.

The first research question of this study is how well-founded Louw's claim is that "text reads text" (Louw, personal communication), meaning that language is its own instrumentation (Louw 2011: 174). This means that it is indeed sufficient to analyze the target text through similar 'events' in the reference corpus, or in the authorial corpus if the analyst is looking for private symbolism, without much recourse to theories and with no recourse to concepts. Theories and concepts are treated by Louw as an unnecessary imposition that obfuscates rather than clarifies. According to Louw, there should be no intermediaries between the researcher and raw data. He says (personal communication): "Concepts are there to explain data; but where data is plentiful, concepts become surplus to requirements, because collocation replaces and supplants intuition and intuitively derived concepts. Collocation becomes the instrumentation that 'concepts' thought they once were in such cases".

Secondly, the study being practical in nature, its other research question is to see how successfully a classroom stylistics methodology founded on Louw's theoretical views can be implemented. My hypothesis is that this methodology may be implemented in a university setting no matter whether the students are native or non-native speakers of the language of the studied text. In this particular case, the students are on a near-CPE[2] level, and their oral performance, for example, may often be described as native. This means that, for a variety of purposes, their feedback may be used to influence the methodology applied in a classroom of native speakers of English. It has been established by Louw as early as in 1993 that even a native speaker's intuition is insufficient to consciously process all the implications of a text (Louw 1993); however, because semantic prosodies (SPs) are frequent, second and foreign language learners also sometimes recognize them. Thus, although the non-native speaker will find more new information in a large corpus than the native speaker, the nature of that information is the same for both. The large reference corpus contains a greater number of encounters with a particular instance of language use than either a native speaker or a non-native speaker has experienced. In the case of a native speaker there may be more cases when the knowledge passes from the known, but not recognized, to the recognized, but the knowledge itself is always the same. That is why the words 'foreign language' are placed in brackets in the title.

However, it is my assumption that, although the principle must remain the same, the level of the tasks should be adjusted to the language level of the students. It is true that the corpus, whether it be a large reference corpus of the language or an authorial corpus, is in fact the sum of our possible encounters with the language/author, but in the case of less proficient students the gap between their personal experience of a pattern and the unedited reference corpus experience will be greater than in the case of their more proficient counterparts.

This study is significant because it puts to the test both theoretical and practical assumptions underlying Louw's CPT. The theory is only sound if it works in practice, and this particular theory insists on being 'instrumentation', on being equated with practice. In practical terms, it is possible for a theory to be sound, but to require a better explanation in the classroom – or a longer one, or more hands-on experience with actual data. This study adopts Louw's "text reads text" approach without much modification, and with the minimum of explanation.

1.2. Methodology

The whole study encompassed both quantitative and qualitative research. The quantitative research consisted of a 'learning phase' and a 'testing phase'. The 'learning phase' was conducted in five sessions. At the beginning of a regular lesson, after a short introduction, the students were given a short excerpt from a text and a concordance, with a particular question to answer. After the answers were written, they proceeded to discuss the text with the teacher and other students. The teacher – myself – gave her interpretation of the concordance lines and the studied text, encouraging the expression of individual opinion. Sometimes a spontaneous discussion of the text and the author's possible meaning ensued. I emphasized that the interpretation is not the teacher's, but should be based on the given concordance and, therefore, their personal analyses of the given concordance were the point of the session. To sum up, the students were learning 'by doing', while being encouraged to express what they saw in the concordance lines and the studied text. This strategy of instruction through practice implemented Louw's stance that a corpus stylistician relies on raw data.

_____

[2] Certificate of Proficiency in English.

Each session of the 'learning phase' contained a different type of task. Each type of task was dealt with once. The students were asked to hand in their answers without making corrections after the discussion and to sign the papers, so that their responses could be marked and the subsequent progress could be monitored. During the first session the main corpus linguistics terminology was introduced and the students had their first encounter with concordance lines. The terms introduced were 'concordance', 'concordance line', 'node', 'collocates', '9-word window' and 'semantic prosody'. The first session only dealt with an authorial corpus (that of Philip Larkin) as it was deemed the easiest type of task. The next four sessions dealt with concordance lines taken from the reference corpus. The reference corpus used for the purpose of the study was the late Tim John's corpus of *The Times* newspaper of the year 1995, containing 44.5 million words. It was originally intended that a session should last up to 15 minutes, but in practice sometimes twice this amount of time turned out to be necessary.

The 'testing phase' was done in one sitting, without warning. The students were given a test of five tasks, mirroring the different types of tasks dealt with during the 'learning phase'. The estimated time of completion was assumed to be 45 minutes, but the subjects were urged to work at their own pace.

Both the results of the 'learning' tests and the final test were processed at the end. A uniform marking scheme had been established for each question. These were the approximate criteria:

- if the analysis fulfils the expectations the mark is 5 (also if the analysis is different from what was expected but excellent and contains detailed argumentation).
- if the analysis of the concordance lines has been done correctly but no connection between the lines and the text has been established, the mark is 4.
- if the lines were incorrectly interpreted, e.g. the student is misled by the first line or mistakenly interprets concordance lines due to a lack of experience, the mark is 3.
- if the analysis is wrong altogether the mark is 2.
- if no analysis was offered at all the mark was 'zero'.

Attached to the test was a questionnaire consisting of 11 questions. It aimed at getting feedback on the short course the subjects had undergone. This was the qualitative part of the research, designed to show what views the students had formed of the text-corpus interaction and of the course. It was conducted to find out if and to what degree the students had taken to the course – whether and how much the students appreciated the course, understood what was going on, found it useful and whether they would choose this subject if they were given the option. It benefited the students as well as the teacher, as it gave them a chance to express their opinions, given that the methodology must have come as a surprise to many of them.

1.3. Background to the study and its limitations

The subjects of the study were second-year students of English at the English Department at the Faculty of Philology, University of Belgrade. These were two groups out of the four comprising the current generation, referred to as group B and group D, and this is how they will be referred to in this study. They represented approximately half of all second-year students. One group was more proficient than the other at the entrance exam, and both may be considered representative of the level of language proficiency of the current generation, as the other half also consisted of one more proficient and one less proficient group. The research was conducted in February and March 2012, in approximately three weeks, within the framework of the second-year *Integrated Skills* course (the whole course was officially named *Contemporary English – G4*), during class time.

As some (though not all) second-year students of English are at the CPE level (some closer to Advanced, others closer to Proficiency), it was thought that their language knowledge was more or less sufficient to attempt the study of concordances taken from reference corpora with a view to interpreting poetic and other texts. None of the students had any previous practical knowledge of either corpus linguistics or stylistics, except for what is taught at secondary schools and on general undergraduate courses in linguistics and theory of literature. By the beginning of their fourth semester the students had completed courses in general linguistics, phonetics, morphology, and had become used to interpreting poetry and prose in their English literature classes. Consequently, their linguistic and academic background was deemed sufficient for them to attempt corpus stylistic interpretation without intervening concepts.

This arrangement had its faults. First of all, not all the students had the same level of language knowledge and yet all were equally tested. It is true that they were all second-year students, but, as stylistics is about nuances of interpretation, the step from CAE[3] to CPE could be crucial and no data of their language knowledge were available except their final results on the *Contemporary English – G3* course (reading, writing, listening, speaking and translation into English and into Serbian) in the previous semester. These results are not wholly reliable, as they partly depended on the students

––––––––––

[3] Certificate of Advanced English.

having learnt certain vocabulary items and grammatical structures (a very proficient student may have avoided the 'cramming' part, and they often do). It is known, however, that, of the two groups, Group B was more proficient on entering the first year of studies than group D. This, the exam results and the teacher's observation suggest that, on the whole, group B was more proficient than group D at the time of the experiment. The teacher's observation during one whole academic year suggests that group B, on the whole, was closer to CPE level and that group D, on average, was closer to CAE. Moreover, the stylistic analysis per se revealed the language level of certain students, particularly in cases where it proved to be lower than desirable. In this context, it was interesting to see whether the group whose level of English was higher and closer to that of the native speaker's would perform better.

Secondly, through force of circumstance, the research was conducted during the *Integrated Skills* course and not in a stylistics course. It was done this way simply because the researcher was currently teaching that course, and had an opportunity of providing instruction and receiving feedback. The main obstacle here was the fact that certain students lacked affinity for stylistic reading of literary texts, found it lacking in motivation and, for this reason, they might not be considered legitimate subjects. It stands to reason that one's performance ought to be motivated if it is to be successful. The students in question may have possessed all the necessary qualities and qualifications and still they may have underperformed through lack of interest. Nevertheless, all subjects were taken into account when the results were being processed. The final qualitative survey was also an attempt to throw light on the issue of the students' interest.

The students' responses varied, so an attempt was made to standardize marking as much as possible in these conditions. There were many variations, so gradations like 3.5 – and even 4.8 – were added. If an analysis exceeded the teacher's expectations in its acuteness, or if the student came up with a correct conclusion or interpretation that even the teacher herself had overlooked, the student was given 6 points – these were special cases. It seemed to me that the difference between a correct interpretation deserving the mark of 5 and an unexpectedly insightful one needed to be documented and taken into account.

No matter how nuanced the marking was, it could not take into account some important differences. First of all, the mark of 4 was given if the student analyzed the concordance correctly, but failed to see the connection between it and the excerpt studied. This included cases where the student was perfectly aware that a connection could be made, but refused to make it, maintaining that the poet meant precisely what he said (during class discussions group D in particular insisted on a poet's freedom not to be 'automated'). Secondly, in practice, since the mark of 2 was given for wrong interpretation, no one was given the mark of 1 – but the mark of 0 existed as part of the marking scheme and was given for no answer. Thirdly, the mark of 0 may have been earned through lack of motivation, as well as inability to offer interpretation. Finally, the teacher's subjectivity in the presence of so many variations is always a threat to standardized marking, despite her conscious efforts to reward similar answers similarly.

2. CONTEXTUAL PROSODIC THEORY TO DATE

2.1. Contextual prosodic theory: main areas of study and literature review

To date, four main areas of Louw's interest are the delexicalization-relexicalization continuum, semantic prosody, subtext and the implications of philosophy of language for collocation. These areas of study are interrelated and have collocation as their pivotal point.

The delexicalization/relexicalization distinction was first brought up by Louw as far back as 1991. This is Sinclair's summary of Louw's idea of delexicalization: "Words can gradually lose their full lexical meaning, and become available for use in contexts where some of that full meaning would be inappropriate; this is the so-called figurative extension" (Sinclair 2004: 198). Relexicalization comes about when a delexicalized word finds itself in the vicinity of a collocate which, purposefully or inadvertently, brings to mind the delexicalized word's literal meaning. For example, in Henry Miller's novel *Tropic of Capricorn*, the words *ghost* and *dead* are part of delexical expressions, but, within the Sinclairan 9-word window they relexicalize (Louw 2006): "Once you have given up the ghost, everything follows with dead certainty, even in the midst of chaos" (Miller 1966: 9).

The idea is explained in Louw (2007, 2008), the latter paper suggesting that "all devices relexicalise" (Louw 2008: 258), and proposing that all devices be given corpora-attested definitions. The 2008 paper also places collocation in the context of Firth's (1957) context of situation, but with emphasis on the provision of corpus-attested terminology.

Semantic prosody, first mentioned in Louw (1993: 157) as the "aura of meaning surrounding a word or phrase", was given a more comprehensive character in Louw (2000), where he first uses the term *Contextual Prosodic Theory*. In the paper he claims that Contextual Prosodic Theory is confirmatory of the Firthian tradition rather than new. A semantic prosody is here defined as

> 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 the speaker or writer towards some pragmatic situation. (Louw 2000: 56; my emphasis)

According to Louw, the key to semantic prosodies is Firth's taxonomy for the context of situation. A semantic prosody arises from a fractured context of situation, and by *fractured* Louw means either under- or overprovided one. The approach falsifies Halliday's grammatical metaphor (Louw, personal communication).

Both these aspects – relexicalization and semantic prosody – are mentioned in Louw (2010a), the very title of the paper suggesting that "collocation is instrumentation for meaning". The paper proposes to dispense with concepts, stating that collocation alone interprets both fact and fiction, and, for the first time, introduces the notion of *subtext*, rooted in the work of analytic philosophers: Frege, Carnap, Wittgenstein and Russel. Co-selection chunks states of affairs, while subtext (quasi-propositional variables) provides the underlying argument, in which the grammatical pattern collocates with the author's lexical choices, falsifying the Vienna Circle's assumption that logic and metaphysics must never be separated. Subtext continues to be studied in Louw's (2010b) examples from Yeats. The application of subtext to prose is illustrated at length in Milojkovic (2013), and to poetry in Louw and Milojkovic (2014).

Milojkovic's contribution to Louw's CPT is the application of semantic prosodies and subtext to Russian (Milojkovic 2011a), pointing to the theory's universality, and the notion of 'grammatical strings' (as opposed to 'lexical items') having an aura of meaning (e.g., *but when did* in Milojkovic 2012). The only claim that a grammatical string may have a prosody, and not using this terminology, was made by John Sinclair in 2006. The grammatical string chosen by him (*when she was*), which was arrived at by choosing the next most frequent collocate, in the end turned out to contain two opposite distinct semantic prosodies, depending on the context of situation (Sinclair 2006). In fact, the two specific fractured prosodies found by Sinclair and embedded in the context of situation prove CPT. My other, very small contribution, is the term 'prosodic clash', describing a situation when a particular writer's or speaker's use of collocation is remarkably different from the prosody established through a reference corpus. It is basically the same as Louw's 'fractured prosody', but Louw's term works within Firth's context of situation, whereas a 'clash' emphasizes the discrepancy between the writer's use and general usage. Within Louw's theoretical framework the term 'prosodic clash' is more telling than 'collocational mismatch', for example. A prosodic clash is an indication of either irony or insincerity in the dichotomy first described in Louw (1993).

Thus, CPT is supported by confirming the thinking of the analytic philosophers and founded entirely on collocation. Co-selection chunks states of affairs in terms of context, produces literary devices in terms of expression, and constantly creates new meaning in terms of semantic prosody (lexical co-occurrence) and logical prosody (subtext). All three need to be viewed on the higher level of events within Firth's context of situation. The implications of CPT are not limited to stylistics, as its terminology may be used to interpret events both fictional and real, by comparing the event in the studied text with similar ones in the huge reference corpus.

To my knowledge, apart from Louw, only one author uses reference corpora in the analysis of (literary) texts, namely, Bettina Fischer-Starcke in her work on Jane Austen (Fischer-Starcke 2010).

2.2. The implications of Contextual Prosodic Theory for classroom stylistics

This section points at those aspects of Louw's theory that are particularly relevant to the present practical study, and will in part draw on examples selected for the subjects' interpretation.

The underlying principle of Louw's stylistics is that "text reads text". It means that no concepts are necessary in order to interpret literary or non-literary texts, but that all we deal with is the reference corpora as the norm against which we judge the text's deviation. Concepts, that is, "the ideas meaning of words", according to Louw (2008: 248, following Firth 1957: 181), are an unnecessary imposition that obfuscates the meaning of a text, while all we need for its successful interpretation is raw data accessible through large reference corpora. This is summarized by the title of one of Louw's papers, "Collocation as instrumentation for meaning" (Louw 2010a), where collocation is seen as a tool which constantly creates meaning through co-occurrence. Situational meaning (meaning in the context of situation and culture) is created in the form of events. A target text contains an event, comparable against similar events in the reference corpus and, therefore, in the world as represented by a balanced and representative reference corpus, created especially for the sampling of the world and creating its dictionary. A line of best fit usually subsists between the target event and those in the reference corpus.

According to Louw, any text can be read against the background of similar texts and the events they represent. Contextual interpretations of keyness are still in their infancy (Louw, personal communication). What is the similarity that qualifies corpus data to act as a background to such a reading? There can be many instances of this, but let us look at a few.

A key word in an authorial text may be 'checked' in the reference corpus to see in what sort of contexts it tends to occur. This analysis may result in revealing a semantic prosody. The author's usage may be clarified by establishing that there is a positive, negative or specific prosody in the language. Alternatively, a clash between the author's use and general usage may improve interpretation in discovering a conscious irony or a subconscious insincerity. In the well-known (to those familiar with the term 'semantic prosody') example of David Lodge's *Small World*, it is said that conference goers are bent on self-improvement (Lodge 1984), whereas in a reference corpus the prosody of *bent on* is

negative and points to destructive intentions. Many semantic prosodies, though perhaps not all, have been mentioned in dictionaries as part of definitions. As native speakers do not consult dictionaries, they may thus remain unaware of certain semantic prosodies in the language, as the example with *cook up* will later show. It is a logical assumption that, where a search in a reference corpus detects a semantic prosody that is used to create irony, irony in the target text cannot be ruled out, as an ironic intention is a definite likelihood.

A similar analysis studying a key word within an authorial corpus may reveal a semantic prosody that is indicative of the author's attitude throughout the corpus of his work, e.g., when Larkin uses the word *day* in prevalently negative contexts and *night* in more positive ones. An absence of a word from an author's corpus may also contribute to the understanding of his work, e.g., Larkin uses the word *night* 72 times, but the word *hope* only nine times. The case of semantic prosody means that collocates of the node in the reference corpus, however diverse, throw light on its usage in the language, if the 'aura' is persistent. In the case of absent collocates, however, there are specific collocates that keep re-appearing. These 'usual', expected collocates may be 'replaced' by an unusual one in the authorial text. For example, in Adrian Henri's poem entitled "Drinking Song" in the line *as the afternoon wore off*, a native speaker will feel that drugs usually wear off and afternoons usually wear on. A non-native speaker will feel this to the extent of his/her proficiency in the language. The non-native speaker will feel this as some kind of word play underlying a metaphor, but a corpus will show what exactly has been replaced. In the quoted line, for example, even a native speaker, when reading, may only notice the abrupt ending of the afternoon, but not the 'drug' part of the pun, which makes it metaphorical. Absent collocates thus contribute to the interpretation of a text and may also help in subtler cases that are not so obvious to the 'naked eye', not as yet armed with corpus experience. It is conceivable that second or foreign language speakers may have been less exposed to the full range of collocates in specific situations. This will mean an increased level of difficulty in dealing with events fractured because of the omission, so they need collocation lists and contexts to unpack them. An increased level of difficulty presents no problem in the 21st-century because of the availability of corpora.

Not only lexical items, but grammatical strings too, may have fairly specific prosodies of their own (Milojkovic 2012). If we look at the case of *but when did*, it becomes crucial to our interpretation of Larkin's notion of love in the following line, which in its context seems to be positive:

> Admitted; and the pain is real.
> **But when did love not try to change**
> **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?

Studying the line starting with *but when did*, we discover that *but when did* introduces angry rhetorical questions, like *But when did a car salesman ever tell you that you had better walk or take a bus?* The semantic prosody is one of futility.

I believe that the angry and rhetorical part in Larkin's line is subconscious, and that it is a variation on the general irony/insincerity division. It is insincere inasmuch as it is not what the person really feels. Larkin *wants to* believe in love's power to change the world and the rhetorical anger is for the most part subconscious. By the way, when asked to judge if Larkin's line in question is or is not optimistic, out of 52 second-year students of English, 37 (71%) replied that it is.

Interestingly, *but when did* found its way into David Lodge's *Small World* too:

> "I have just one question," said Philip Swallow. "It is this: what, with the greatest respect, is the point of our discussing your paper if, according to your own theory, we should not be discussing what you actually *said* at all, but discussing some imperfect memory or subjective interpretation of what you said?"
> "There is no point," said Morris Zapp blithely. "If by point you mean the hope of arriving at some certain truth. **But when did** you ever discover *that* in a question-and-discussion session? Be honest, have you ever been to a lecture or seminar at the end of which you could have found two people present who could agree on the simplest precis of what had been said?"
> "Then what in God's name *is* the point of it all?" cried Philip Swallow, throwing his hands into the air.
> "The point, of course, is to uphold the institution of academic literary studies."
>
> (Lodge, *Small World*, my emphasis in bold)

Therefore, this classroom corpus stylistics research focuses on semantic prosodies (whether they surround a word, a phrase or a grammatical string), which may or may not be consciously felt by native speakers. They also may or may not have been fully captured by dictionaries, which is a point worth dwelling on. Apparently, there are two factors which might restrict the elaboration of semantic prosodies in dictionaries. The first one is obviously space. The second has to do with the capacity of a definition to fit in with any instance of a word's or expression's use. If we suggest that

but then, or *but then again* in the majority of cases has a narrower meaning to it than expressing a slight contradiction or lack of surprise (COBUILD), or even than lessening the meaning of a previous sentence or lack of surprise (Longman), the explanation must fit every text we encounter. But it is in the nature of a prosody not to extend its aura to every single case. Also, thorough research would need to be carried out on *but then* and *but then again*, involving the study of contexts of situation of specific sentences, which might justify devoting a whole paper to it. Naturally, it may be impracticable when compiling a dictionary, although perhaps not inconceivable in the future. For example, *but then* and *but still* are different, and not only in the degree of contradiction. Since creating a dictionary involving all nuances of meaning is practically impossible, that is where corpus stylistics proves useful in relation to a text's interpretation. It is hypothetically possible that the aura of the comforting *cest la vie* meaning of *but then* is extended to the remaining 32% not carrying it at first sight, if one studied the contexts of situation for the remaining 32%. This needs to be researched in the future, and will hinge upon discovering collocates that differentiate these two senses. A question of what is and what is not to be found in corpus-based dictionaries and why could come from students studying a given concordance. For example, some sceptical and conservative subjects of this study were surprised to learn that the negative prosody of *bent on* is recorded in the COBUILD dictionary.

The other methodological question that bears upon lexicography, corpus stylistics and corpus stylistics pedagogy has to do with the amount of percentage of occurrences of a specific tendency in meaning. How far does the aura have to spread for the word or phrase to be considered as bearing a prosody? Is 60% enough? 65%? 75%? How much is enough for a lexicographer? How much for a stylistician? What do we tell the students in the classroom who attempt to interpret texts? Do we tell them to rely on corpus-based dictionaries for prosodies in order to avoid error?

The answer to these questions has to do with the term 'events' (Louw and Milojkovic 2014). The mere proportion of lines showing a prosody is not sufficient proof that a text should be interpreted in a certain way. In the tradition of the philosophy of language and its postulates (Russel 1948) and according to the notion that a reference corpus is a representation of the world and its dictionary, we must look for *similar* events. The lines that carry the investigated pattern, whether based on a grammatical string or a lexical collocation, and describe similar events are the ones to be considered. This advanced notion was not shared with the second-year students, who were instead simply at the mercy of concordance lines. However, 10% of them did mention that in a concordance certain lines corresponded in meaning to the one in the studied text, which must have involved their studying contextual clues in these concordance lines. This finding shows that, at least to these students, text does read text, as they could see the notion of events without previous instruction.

3. QUANTITATIVE RESEARCH

3.1. The learning phase

This section of the paper will describe the tasks given to students during the 'learning phase' of the experiment, the instructions they were given and their response. There were five learning sessions on the whole, one task per session. The sessions took place at the beginning of regular *Integrated Skills* classes.

At the very first encounter the students were given full concordances from the authorial corpus of Philip Larkin with the nodes *day*, *night*, *light* and *God*.[4] As Larkin's use of these words differs drastically from the conventional – *day* is viewed pessimistically, *night* brings relief, *light* appears as dark, and *God* is doubted –, this was a good opportunity to illustrate to the students 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 to facilitate comprehension at this first encounter. For example, this is the concordance with *God* as the node:

---

[4] My paper on Philip Larkin (Milojkovic 2011b) is available online, with the four concordances given in full: <http://www.belgrade.bells.fil.bg.ac.rs/Bells 3.pdf>.

```
                    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
                Data from the following files: ZARKIN.CTX
```

Figure 1. Concordance for *God* as node

At the second encounter, the students were given the excerpt from David Lodge's novel *Small World* quoted in Louw (1993). Unlike the first introductory session, the second one involved a task that was to be completed individually in a test-like format before the feedback and the 'right' answers were given. In the task, the students were first asked several comprehension questions, and then given chosen concordances from *The Times* corpus, printed in a bigger size than usual and with the significant collocates in bold. The concordance had been edited to facilitate understanding, without the students' knowledge, as the unedited concordance might have discouraged them. In this session, the students were taught the notion of semantic prosodies and explained that the prosody of *bent on*, which they had just discovered, is in the dictionaries. After it was elicited from the class that Lodge is in fact being ironic, Louw's irony/insincerity dichotomy was explained to the students. This is the format of the task:

Session 2 (Semantic Prosody)

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 ☺

3. 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?

Figure 2. Session 2 (semantic prosody)

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. The task proved a success not only because the students could see the role played by absent collocates, but also because learning phrasal verbs is always of interest to non-native speakers. Whereas during the previous session the subjects reacted with surprise to the notion of semantic prosody, this time they

understood the practical advantage of establishing the meaning of a word through its collocates in the reference corpus. This is the format of the task, whose impact is briefly discussed in Section 2.2:

```
Session 3 (absent collocates)

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 exquisite
 2 y to stare at him. The novelty soon wore off, however, as Smythe persistently re
 3 lbeing that it was months before it wore off. I am still trying to remember wh
 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 explain
 6 men using implants said the effects wore off more quickly, and 29% said they nee
 7 ut eventually, inevitably, the drug wore off. Some say it was Fortensky who call
 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
                   Data from the following files: TIMES95.TXT

                   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 not
 3 only obscured my face but, as time wore on,had a horribly isolating effect on m
 4 things were to change as this game wore on. After 17 minutes, Durrant put the
 5 on to the bat and off it as the day wore on. Although he found life more diffi
 6 officials reported that as the day wore on an ever-growing crowd of terrified o
 7 Corsie's form improved as the match wore on, and to whitewash a player of Schuba
 8 w more perfunctory as the afternoon wore on and finally ended up with Stewart ha
 9 was only sustained, as the evening wore on and got colder, by the particular in
10 tive electoral history. The night wore on and the flow of Tory setbacks mounte
11 did get harder, and time certainly wore on. And on... But first, Wimbledon. T
12 d dummy runs got better as the game wore on and some of his early ones were no m
13 had become even firmer as the match wore on and were not going to be pulled arou
14 ut he played ever better as the day wore on and, after Ilott had returned to dis
                   Data from the following files: TIMES95.TXT

Comment:
```

Figure 3. Session 3 (absent collocates)

The fourth session focused on grammatical strings and their prosodies. The subjects were first asked whether they perceived the lines containing the grammatical string under study as positive or negative. Then, they were given the whole concordance from *The Times* corpus together with the wider contexts of four chosen concordance lines. The wider contexts were provided in order to better illustrate the relationship that is usually established between the clause starting with *but when did* and the surrounding text. The wider contexts of lines 1, 2, 3 and 6 were chosen because they were easier to understand than others, since newspaper language with its ironies, sarcasm and sophisticated vocabulary is not always easy for non-native speakers. This is the format of the paper:

Philip Larkin
  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
  The world back to itself – no cost,

Is this an optimistic reference to love? 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 that
 2 s may make little economic sense, but when did economics really come into the eq
 3 yd and Rob Lowe also participate, but when did either last make a prudential car
 4 ties in both manager and country. But when did England last have success or a pu
  5 be the logical time to bow out. But when did football, life, logic, Charlton o
   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 why
 8 company, making £2billion a year. But when did you last hear critics sounding of
  9 hormone is, of course, a cop-out, but when did you last hear of a netball crowd
                    Data from the following files: TIMES95.TXT
```

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?*

4. Two related programmes on BBC2 focus on elephants and their would-be savior, Richard Leakey. *The Savage Paradise* (Monday, BBC2, 8pm) is billed as an intimate portrait of an elephant herd in Botswana, while Leakey is profiled in *Africa's Wildlife Warrior* (Wednesday 9.30 pm).
  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 Warrior 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?*

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?

Figure 4. Session 4

   As mentioned earlier, on the first reading, 71% of the subjects saw the reference to love in the poem 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 the concordance lines in the reference corpus, especially as it was a newspaper one, had nothing to do whatsoever with the poem and the poet. The poet, they claimed, was free to use the language as he pleased. My 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 my hypothesis that the poet intended to make a positive statement while subconsciously he did not believe in the power of love to change the world back to itself.

The fifth section of the experiment will not be described here in detail as it was unsuccessful. During the fifth section 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 well 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, quoted in 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. Had the experiment been conducted as part of a stylistics course, I would have been justified in spending more time on teaching subtext to students. As it was taking up the time of an altogether different course, with different aims and a set curriculum, I abandoned subtext until a better opportunity arose. This does not disprove the principle of "text reading text", but rather calls for more time spent on explanation and classroom practice.

3.2. The testing phase

All tasks in the first four sections were mirrored in the final test. In the first task, the students were given an edited concordance with the node *hope* from the authorial corpus of Philip Larkin. The second task had to do with a negative semantic prosody of the phrasal verb *cook up*, the prosody not being obvious solely on the basis of the context. The text itself came from Leo Jones' *New Progress to Proficiency* (2001: 112) and had recently been studied in the classroom, so the subjects were familiar with the wider context as well, although this 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 analyze wider contexts from *The Times* corpus, and in the fourth the grammatical string needed to be interpreted on the basis of a concordance.[5] The fifth task was based on the third step in the learning phase and had to do with 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 analysis of the concordance and asked to connect it to the studied excerpt from another poem by Larkin. Here is the format:

---

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 nine 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 somethi
  5 what I desired - That long and sickly hope, someday to be As she is - gave a fli
  7 e Through doubt from endless love and hope To hate and terror; Each in their dou
  8 it's a different country. All we can hope to leave them now is money. 10 lanuar
                  Data from the following files: ZARKIN.TXT
```

Comment:
1. 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 da
  2 h of Euro-scepticism, and you have cooked up a crisis." Tory Euro-sceptics wil
  3 a stream of mixed notices, having cooked up a storm in America. "Crime in exces
  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, whic
 11 liance claimed the affair had been cooked up by the Russians in an attempt to de
 13 e fallen for every publicity stunt cooked up by the lawyers in the Simpson case
```

---

[5] All three concordances in the final test were edited – given that the time was limited – and some unmotivating lines were removed, but the necessary level of difficulty was preserved.

```
17 ese than the Mayan extravaganza he cooked up. Certainly the claim that the build
18 mmary of the predicament Slovo has cooked up for her headstrong part-time detect
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 c
23 mmon murderer. Salvatore Cammarano cooked up the plot and later provided somethi
25 , 20, 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 councillors
29 ant turns out to be less than it's cooked up to be, and Connie's disillusionment
30 ed up the National Lottery (I said cooked up, you sniggering lot) and departed b
                   Data from the following files: TIMES95.TXT
```

2. 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 *The Times* corpus:

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.
Ibsen himself was subject to fits of depression, so he wasn't one for light entertainment. *But then*, few Norwegian entertainers are.
'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.'
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?
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.
'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 derived from the *The Times* contexts.

3. Now read the contexs of *but what is* found in the *The Times* 1995 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, rec
 3 r desk pontificating arrogantly". But what is a columnist for if not to pontific
    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 book
 42 are ball into the six-yard box. But what is he meant to do, other than act as
43 preference holders. Clear enough; but what is he doing upping his stake in anoth
 56 aling, though possibly necessary. But what is it all for? "Have some knowledge
58 ibition spaces in central London, but what is it beyond that? The Academicians t
```

```
 73 r remains far removed from normal but what is normal behaviour for a king? Seizi
 79 oes not look like a comic genius, but what is one supposed to look like? As he l
 80 enstern, all we get is incidents; but what is our role in these incidents? Have
 95 claims are "exorbitant demands". But what is reasonable? Sybil Gooldrich, one
 97 rought down to earth with a bump. But what is risk, and how can you avoid it? Th
112 deaf members of the audience." But what is the point of interpreting opera fo
114 ood as Claridge's or at my house, but what is? The clientele were an odd mix,
116 ould love to lead the Government, but what is the point if the party is too asha
117 he accused sold were not genuine. But what is the difference between a genuine l
118 ng as many sights as time allows. But what is the rush? Rome was not built in a
119 ow inflation and a trade surplus. But what is the point? This is a question whic
127 s is impressive and he is excited but what is the reality? Market forces in ed
130 dangerous Rollerblades can be. But what is the attraction? Unlike traditional
139 ically/she's using him fiscally." But what is this thing called friendship? When
144 ays before an international game. But what is to stop the clubs refusing to sign
                    Data from the following files: TIMES95.TXT
```

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.

4. And everywhere the stifling mass of night
    Swamps the bright nervous day and puts it out.

The 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, and 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. Four concordance lines contained *put it out*. In all the four 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?

Figure 5. Final test

3.3. Discussion of results in the quantitative phase

All the students' tests, in both the learning and the testing phases, 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. Column '1a' is empty because the subjects' answers were not graded during the first session. Column 4a is empty as task 4b is focused on a grammatical string and as well as 3b.

These results were processed in the following way. The scores of both groups were entered into separate tables. The students were given marks from 0 to 5. Grade 0 was given to students who were present, but left a blank space instead of doing the task. In practice, since a completely wrong answer was given the mark of 2, no one was given the mark of 1. In cases of exceptionally acute judgement the mark was 6 out of 5. The average results per each task were calculated, for the two groups separately as well as for all subjects together. The average marks were calculated per each group and for all the subjects both in points and percentages (the mark 5 points for all the students was accepted as 100%). Statistical treatment was performed in an *Excel* spreadsheet. The final marks were plotted in a diagram, for the two groups separately, as well as for all the students together.

Group B	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	sum	%
Average		5	5	4.2	4	4		4.3	4.6	4.1	21.38	85.509
Standard deviation		0.7	0	0.7	1	1		1	0.4	1.1	2.20	8.8095
Number of students present		22	9	22	20	22		22	18	22	22	22

Table 1. The results of Group B

Group D	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	sum	%
Average		4.5	5	4.1	3	3		4.1	4	3.7	19.68	78.714
Standard deviation		1.1	0	0.8	1.3	1		1.4	1	1.8	3.246	12.982
Number of students present		28	10	28	24	28		28	23	28	28	28

Table 2. The results of Group D

Group B + D	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b	sum	%
Average		4.7	5	4.14	3.33	3.58		4.14	4.28	3.87	20.4	81.7
Standard deviation		0.95	0	0.76	1.28	0.86		1.22	0.86	1.51	2.93	11.74
Number of students present		50	19	50	44	50		50	41	50	50	50

Table 3. The results of the two groups together

The difference in the marks of the two groups for the 'b' tasks is statistically significant (p=0.041). It is obvious that the subjects from group D scored fewer points than those from group B. Figure 6 shows the plotted curves of the distribution of marks for both groups and for all students together.
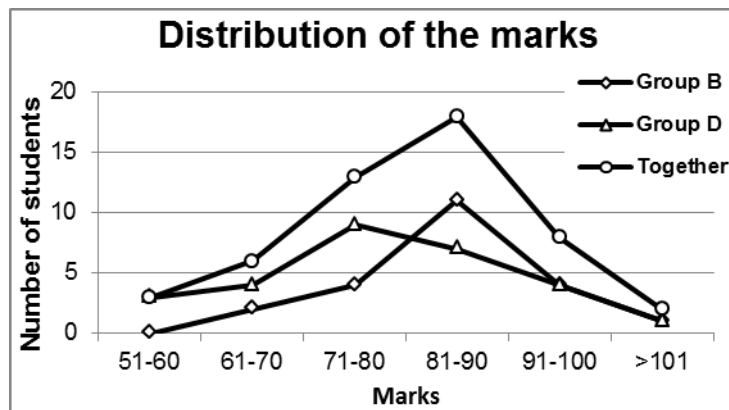

Figure 6. The distribution of final marks

Although the results scored in group D are lower than in group B, which confirms the initial assumption that the more proficient group would score better results, the results of both groups suggest that the difficulty of the tasks was adequate, since all three curves could have been obtained after any undergraduate course of moderate difficulty.

As the first research question is to see whether "text reads text" for Belgrade students of English, it is important to establish the percentage of students who 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 has been 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 our research question than the previously given 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 which is relevant to the research question is that five students out of fifty (10%) made comments when studying concordance lines in 4b (*but what is*) that could be construed as attempts to look for similar events in the reference corpus. As such a method of interpretation had not been mentioned in the classroom, this also confirms Louw's stance that "text reads text".

4. QUALITATIVE RESEARCH

Attached to the final test was a questionnaire consisting of 11 questions whose aim was to see how well the students understood the point of the course, whether they found it useful, what they thought of the methodology, whether they enjoyed it and whether they would choose it if it were on offer. Here is the format of the questionnaire:[6]

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?

Figure 7. Final questionnaire

The first three questions were aimed at discovering what definitions of 'corpus linguistics', 'corpus stylistics' and 'stylistics' the students would give after being exposed to a short course that, in itself, was based on the principle that "text reads text". The researcher was curious to see how the subjects understood these three disciplines. 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'.

Out of 50 students, 52% gave acceptable definitions of 'corpus linguistics'; 48% defined 'corpus stylistics' (6% more defined it by means of the word *style*) and 32% defined 'stylistics' (16% more defined it by means of the word *style*). I have separated the answers which depended on the word *style* as it 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 ones with a 'no'. To preserve this principle of describing answers as 'yes' if the feedback is positive and 'no' if negative, in question 10 'yes*' denotes that the course was not found difficult by the student, and 'no*' means that it was found difficult.

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 course were not marked with a 'yes'). The adopted teaching methodology was approved of by 72% of the subjects, and 48% stated that the course was not difficult.

In the last question, consisting in fact of two, one related to the enjoyment of the course and the other to whether the student would choose it if it were 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: 'yes/yes': 36%'; 'yes/maybe': 16%; 'yes/no': 22%; 'no/no': 14%; 'no/yes': 4%; and no answer: 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 they had enjoyed it. These percentages would 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 paper proved to be successful.

5. CONCLUSION

The two interrelated research questions posed in this study are a) if "text reads text", i.e., if reference corpora alone and without theoretical concepts can help interpret authorial text, and b) if the proposed CPT-based classroom stylistics

---

[6] The questionnaire was in part inspired by Burke (2004).

methodology can be successful. After the quantitative and qualitative survey conducted in this study, there is sufficient foundation for the claim that both research questions have been answered positively. The percentage of 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. The fact that the more proficient group achieved noticeably better results suggests that native speakers would have been even more successful, but that the principle is the same regardless of the level of proficiency. The implication is that the degree of tasks' difficulty may – or should – vary depending on the students' or pupils' proficiency. The feedback gained on the qualitative part confirmed the success of the teaching methodology. Both quantitative and qualitative results in fact exceeded the researcher's expectations, given the study's limitations, which deserve some dwelling upon.

Firstly, had only the motivated students been tested, the final results might have been even more encouraging. As things stood, a fair amount of students were not particularly interested in poetry, corpora or stylistics. Another issue is the level of personal, and not linguistic, maturity, which will be reflected in a stylistic interpretation of the poem or of the lines. Personal maturity in the issues of love, hope, despair or resignation is a factor altogether different from, for example, the inability or refusal to see that a reference corpus does have a bearing on a particular author's meaning. The subjects, in this case young adults, may not have the experience that the (middle-aged) author has tried to convey and, therefore, cannot interpret what they have not understood. When interpreting a poem or part of it through concordance lines, it is first necessary to check each student's literal understanding and the degree of their appreciation of the text as readers. This 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.

It is also worth noting that no proper course of corpus stylistics would have been founded on such a minimum of instruction. Therefore, the point of this research, again, is to see how the subjects react to texts. However, in real life, more students would have responded to this kind of teaching positively after reading on semantic prosody. Some students may have understood the point, but were too conservative to believe it, as it may not have fitted into the manner of their dealing with text and meaning in their previous schooling. They could have changed their minds after reading a couple of papers containing examples. As things stand, some of them may have been too conservative to start seeing things differently. For example, one of the students, generally proficient and hard-working, could only see grammatical usage in the lines, but never auras of meaning. On the other hand, they might have complied with the expectations of a regular course unquestioningly, whereas the adopted way of learning showed how they felt when not under pressure.

It is obvious from the above that the marks should not be interpreted only as comments on the quality of the students' answers, but rather as a way of showing what kind of feedback a subject gave. A difference must be made between a student who is not capable of perceiving how a concordance can assist the interpretation of the studied text 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 studied text had not been made – whether the student was not able to make the connection or refused to make it.

Another restriction has to do with the fact that the final test was slightly more difficult than the 'learning' tests, 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, namely, 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 studied line in the text better. Ironically, in some cases they claimed it even if objectively they misunderstood the line.

A lack of basic skills in reading concordances on the part of the subjects was another important limitation. Overgeneralization was one of the observed errors – sometimes 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, namely, they were used to the sentence, clause or syntagm as units of interpretation, rather than a concordance line that could begin or end at any point in a sentence, clause or phrase. All these issues would have been addressed on a proper corpus stylistics course.

With hindsight, my general impression is that the final test might have been too difficult for a fair number of my students, due not exactly to lack of linguistic proficiency, but 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 turn out in a real-life situation, 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.

REFERENCES

Burke, Michael. 2004. Cognitive stylistics in the classroom. *Style* 38/4: 491–510.
COBUILD. 1998. *Collins COBUILD English dictionary*. London: HarperCollins.
Jones, Leo. 2001. *New progress to proficiency. Student's book*. Cambridge: Cambridge University Press.
Firth, John R. 1957. *Papers in Linguistics 1934–1951*. Oxford: Oxford University Press.
Fischer-Starcke, Betina. 2010. *Corpus linguistics in literary analysis: Jane Austen and her contemporaries*. London: Continuum.
Henri, Adrian, Roger McGough and Brian Patten. 1967. *The Mersey sound*. First edition. London: Penguin.
Larkin, Philip. 1988. *The collected poems of Philip Larkin*. London: Faber and Faber.
Lodge, David. 1984. *Small World*. Harmondsworth: Penguin.
Louw, William E. 1991. Classroom concordancing of delexical forms and the case for integrating language and literature. *English Language Research Journal* 4: 151–178.
Louw, William E. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Mona Baker, Gill Francis and Elena Tognini-Bonelli (eds.), *Text and technology: in honour of John Sinclair*. Amsterdam: John Benjamins, 157–176.
Louw, William E. 2000. Contextual Prosodic Theory: bringing semantic prosodies to life. In Chris Heffer and Helen Sauntson (eds.), *Words in context. In honour of John Sinclair*. Birmingham: ELR, 48–94.
Louw, William E. 2006. Collocation as the determinant of verbal art. In Patrick D. Miller and Monica Turci (eds.), *Language and verbal art revisited: linguistic approaches to the study of literature*. London: Equinox, 149–180.
Louw, William E. 2007. Truth, literary worlds and devices as collocation. Closing keynote presentation at TaLC6 on 7th July 2004. In Encarnación Hidalgo Tenorio, Luis Quereda Rodríguez-Navarro and Juan Santana Lario (eds.), *Proceedings of the Sixth Conference on Teaching and Language Corpora*. Amsterdam: Rodopi, 329-362.
Louw, William E. 2008. Consolidating empirical method in data-assisted stylistics: towards a corpus-attested glossary of literary terms. In Sonia Zyngier, Marisa Bortolussi, Anna Chesnokova and Jan Auracher (eds.), *Directions in empirical literary studies. In Honour of Willie van Peer*. Amsterdam: John Benjamins, 243–264.
Louw, William E. 2010a. Collocation as instrumentation for meaning: a scientific fact. In Willie van Peer, Sonia Zyngier and Vander Viana (eds.), *Literary education and digital learning*. Hershey, PA: IGI Global, 79–101.
Louw, William E. 2010b. Automating the extraction of literary worlds and their subtexts from the poetry of William Butler Yeats. In Marta Falces Sierra, Encarnación Hidalgo Tenorio, Juan Santana Lario and Salvador Valera Hernández (eds.), *Para, por y sobre Luis Quereda*. Granada: Universidad de Granada, 635–657.
Louw, William E. 2011. Philosophical and literary concerns in corpus linguistics. In Vander Viana, Sonia Zyngier and Geoff Barnbrook (eds.), *Perspectives on corpus linguistics*. Amsterdam: John Benjamins, 171–196.
Louw, William E. and Marija Milojkovic. 2014. Semantic prosody. In Peter Stockwell and Sarah Whiteley (eds.), *The Cambridge handbook of stylistics*. Cambridge: Cambridge University Press, 263–280.
Miller, Henry. 1966. *Tropic of Capricorn*. London: Panther.
Milojkovic, Marija. 2011a. Semantic prosody and subtext as universal, collocation-based instrumentation for meaning and literary worlds. In Victor P. Zakharov (ed.), *Труды международной конференции Корпусная лингвистика*. St Petersburg: St Petersburg State University, 47–52.
Milojkovic, Marija. 2011b. Quenched light, or seeing through a glass darkly – a collocation-based view of Larkin's atheism and depression. In Zoran Paunović (ed.), *Belgrade English language and literature studies*. Volume III. Belgrade: University of Belgrade, 127–144.
Milojkovic, Marija. 2012. Time and transitions in Larkin's poetry. *NAWA: Journal of Language and Communication* 6/1: 102–126.
Milojkovic Marija. 2013. 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.
Russell, Bertrand. 1948. *Human knowledge: its scope and limits*. London: Routledge.
Sinclair, John M. 2004. *Trust the text*. London: Routledge.
Sinclair, John M. 2006. *Phrasebite*. Pescia: TWC.