

TEACHING CRITICAL DISCOURSE ANALYSIS OF SOCIOCOGNITIVE APPROACH TO EFL STUDENTS THROUGH CORPUS BASED ANALYSIS

Yogi Setia Samsi*

Universitas Pendidikan Indonesia
yogi.setiasamsi@staff.unsika.ac.id

Iwa Lukmana

Universitas Pendidikan Indonesia
Iwa.di.bandung@gmail.com

Dadang Sudana

Universitas Pendidikan Indonesia
dsudana@gmail.com

doi: 10.35706/eltinf.v3i1.3760

To cite this article: Samsi, Y. S. (2020). Teaching critical discourse analysis of sociocognitive approach to EFL students through corpus based analysis. *ELT in Focus*, 3(1), 37-45. doi: 10.35706/eltinf.v3i1.3760

Abstract

Many studies have been researching critical discourse analysis, henceforth CDA, in analyzing numerous contexts. Nevertheless, those studies still lack of contributions to, for example, pedagogical aspects as well as incorporations of specialized technology such as corpus tool. This study thus intends to fill that gap that is qualitatively aimed to explore ways of teaching CDA focusing on sociocognitive approach through corpus analysis and to know learners' responses towards their learning experience. This study also concerns on three elements: corpus based approach, CDA approach, and pedagogical approach. The target participants were 7th semester undergraduate students of a university in Karawang, Indonesia. They have been studying linguistics and using learning media. The study spanned for one month and included two training sessions for learners on how to use concordance software in analyzing words, phrases, concordance line, frequencies, and collocations which were approached sociocognitively within special microstructure couched by Van Dijk (2008). The corpus were comprised of two selected newspapers with a specific issue on big data. The results showed that the participants are able to use a concordance software independently as it seemed exciting to the students when doing the exercise. It indicated successful way that students eventually can criticize some discourse through corpus analysis. Further study is required to deepen the analysis and incorporate other approaches such as systemic functional linguistics to strengthen the analysis.

Keywords: *Teaching CDA, Corpus linguistics, Concordance, News discourse, Microstructure*

INTRODUCTION

In this development digital era, teaching and learning should always create an innovation especially in delivering the materials to the students. In the reality, the teachers are still confused to make an innovation in their teaching method. It hardly developed by teachers in their

teaching due to lack of insight regarding the innovative and creative way of any courses in classroom. It thus caused the students boring and the learning is monotonous although literature studies many ways in teaching to make students more excited and avoid a tedious teaching. However, this is the essential stuff to be overcome to learning and teaching in classroom. One of

the ways to innovate the teaching and learning is involving the technology. Precisely, by optimizing the information and the technologies have become part of supporting media in the classroom activity. It is such useful and greatly influence to the learning process for both of teacher and students by using computer-based tools as there are many previous researches have been developed over the last decades. Regarding the corpus linguistics, it is also one of improvement tools in digitalized age although it is a surface identification of the actual and real production of language (either spoken or written). The productions of language can be spontaneously taken from various authentic sources and fields such as newspaper, magazine, people's speech and conversation and etc. Corpus-based studies have traditionally been less concerned with whole texts or with the social context and have thus been characterized as working from the 'bottom up' (Swales: 2002). In addition, Biber (1988; 2006) argued that corpus research has played a key role in distinguishing the overall characteristic of academic prose by means of multi-dimensional analysis.

From the aforementioned information, it can be used to investigate the material to get either the formulation or any ideologies in discourse analysis perspective. It is also as the gap amount previous research while other previous study is strongly expressed about method, and grammar itself. In teaching case, one of them studies on Data-driven learning for teaching collocations of learner performance, proficiency, and perceptions (Vyatkina 2016). In addition, DDL is not only related to English learning and teaching but also to the core of linguistics, extended linguistics, and hybrid linguistics. For instance, Yu Hou (2014) found that corpus linguistics is used to identify nominalization in translation of Chinese literary prose. Furthermore, Kim and Chun (2008) study more focused on lexis awareness through corpus based data-

driven learning. However, the study on a specialized critical discourse analysis, henceforth CDA, rarely found and lack of incorporations with other studies. Practically, it seldom demonstrates concordance application to analyze the text to get the dominant of speech, the power of language, or to know hidden ideology. Another is Adel and Reppen (2008:2) arguing for 'the viability of corpus-based research and corpus-assisted tools for discourse studies'. Further is from Ken Hyland in Chales's book (2009) about corpus informed discourse analysis: the case of academic engagement. The mentioned studies are still too general through discourse analysis, but nevertheless, the study on specific to integration of CDA and teaching aspect is far from the literature.

Therefore, in this present study, it enthusiastically aimed to investigate teaching linguistics interfaced with corpus linguistics such as teaching critical discourse analysis. The objective of the study is focusing on how to teach CDA following Van Dijk (2008) of sociocognitive approach through concordance software and students' responses of their experience. Considering the previous studies above, as far as the literature views that this kind of study is still rare to conduct in Indonesia. It should be noted that this present study concerns into the implementation of a specialized corpus on teaching CDA to analyze sociocognitive approach that specified by only micro structure in order to analyze the authorial media represented the ideology in news articles and the influences into society perspective.

Corpus Linguistics works

McEnery et.al (2006:7) argue that corpus linguistics has gone 'well beyond [its] methodological role' and has become an independent 'discipline'. It can be optimized by using computer software, it is called concordance software. Briefly, I

outline some techniques or corpus processes that can be carried out on corpus data as Baker (2010). *First*, frequency, it is the bedrock of corpus linguistics. At its simplest level, frequency refers to the numbers of times something occurs in a corpus (or text). Frequency counts need not to be limited to single words. It is possible to calculate frequency of grammatical, semantic, or other categories. *Second*, collocation, identified by Firth (1957), is a way of demonstrating (relatively) exclusive of frequent relationships between words (or other linguistic phenomena). If two words collocate, then they have a tendency to occur near or next to each other in naturally occurring language use. For example, *tell* and *story* are collocates because they occur in a range of different grammatical contexts such as *tell me a story*, *story to tell*, *let the story tell itself*, *tell a story*, and *that story does not tell us anything*.

Third, keywords are a way of taking into account relative frequencies between corpora, which is a useful way of highlighting lexical *saliency*. For example, the word *the* is generally very frequent in most corpora, so knowing that it is frequent in a corpus that we are examining may not be particularly exciting – is simply tell us that our corpus is typical of most language use. *Fourth*, a concordance is a table of all the occurrences of a linguistic item in a corpus, presented within their linguistic context (usually a few words to a few lines either side of the linguistic item). Concordances are an important aspect of corpus linguistics in that they allow qualitative analysis to be carried out on corpus data, letting the researcher explore individual cases in detail. Sorting concordance data alphabetically is an often-used way to identify patterns quickly and also on a different word position is likely to produce different patterns. Simply, concordances also allow the researchers to identify linguistic patterns, which can be based on grammar, meaning, pragmatics, and discourse.

Sociocognitive Approach

Social cognition approach developed by Teun A. Van Dijk (2008) that focuses on issues such as ethnicity, racism, and refugees. This approach is referred to as social cognition, because he sees cognition factor as an important element in the production of discourse. Therefore, this approach discourse analysis can be used to determine the social position of ruling groups or dominant and marginalized groups. Further, he assumed in Wodak (2009) that discourse analysis is not limited to the structure of the text because the structure of discourse itself indicate or signify a number of meanings, opinions and ideology. On the other hand, in order to reveal the hidden meaning of the text, it should take the analysis of cognition and social context as sociocognitive. He divided into three levels of textual analysis, namely; (1) micro structure, (2) super structure, and (3) macro structure. In the micro structure, Van Dijk highly concerned to the theme and rheme in the text as textual meaning grounded by Halliday (2004) as the focus of this present study. Further, he also differed the super structures into (a) summary that contained *title* and *lead* in the text, while (b) story divided into *situation* and *commentator*.

The last is micro structure that classified into several components such as transitivity system, nominalization, passivation, and references. From the previous study aforementioned, there is a specific area that has not found in other literatures. It is still seldom to focus on interfacing teaching CDA grounded by Van Dijk (2008) and corpus based study. Therefore, it should be noted that this present study specifically discussed about only micro structure based on sociocognitive theory in this gap.

METHOD

In this part, the method is clarified that this study was qualitative-descriptive.

Moreover, Ary, Jacobs, Sorensen & Razavieh (2010) point out mentioned research design aims to describe phenomenon and to reveal subjects' perspective on what they experienced. The participants involved 20 students of 7th semester of English education department in one of university in Karawang, Indonesia as they were sincerely gathered to participate as learners in this study. Further, the corpus for the study was compiled from news article collections about government regarding Rohingya's issue in BBC news as International newspapers from January-May in 2020 consisted of twenty articles about 200.000 words. The compilation corpora are attached into concordance software to category, identify, and interpret the words, phrases, and contexts based on frequency, collocation, and concordance line. The software enables to gain a certain news text that contains with a specific term such as typing a word or phrase in the search of keyword in context, hereafter KWIC, for instance: *government**, OR *Trump**, etc.

In the first step of study was the teacher explained the origin of sociocognitive approach limited only microstructure in CDA to the students where the corpus software had already served too in students' computer. Secondly, in the training sessions, the teacher applied the software of concordance to demonstrate to students for a first step of sociocognitive identification. The teacher opened the software and type as KWIC such as wording; *government**, *Trump**, or nominalization; **ing*, **ment*, **ion*, **ation*, and **ed*, etc. After typing, the appearance word was directly categorized, analyzed, and interpreted based on the CDA theory in order to get the ideology of the writer from the issues. However, those procedures were demonstrated in learning and teaching to know the divergent between teaching CDA by using software and conventional one. It was the way where the teacher demonstrated the students to teach CDA

and found out the appropriate word to reveal politic issue through KWIC while displayed also its frequency, collocation, and context. In the last step of study, the teacher instructed and worked together with the students to discuss the result and their relevance with the theory. Therefore, subjectivity is a needed to criticize the content of the searched word contextually until unpacking the implicit meaning or revealing the hidden actor/writer inside of the text. Eventually, to strengthen the result and answer the further aim, the researcher takes questionnaire in order to obtain the students' responses toward the implementation of learning CDA and using the corpus tool as they experienced.

RESULTS AND DISCUSSION

To identify sociocognitive through Van Dijk (2008) analysis framework by finding out macro structure, super structure, and micro structure. As limitation in this study, it is focused on micro structure that covers (1) transitivity system, (2) nominalization, (3) passivation, and (4) reference. Those analysis of aforementioned elements are covered in the discussion of teaching CDA through corpus software as the first sub-theme in this part and the second part cover the discussion of corpus software implementation in teaching CDA. Those discussion will be further delineated as follows.

Teaching CDA through corpus software

In this part, the description of the study combines the result of implementation of corpus analysis in teaching CDA in classroom and the description of microstructure analysis based on corpora. It expressed the activities of learning and teaching including training in the classroom as follows.

In the first analysis, it found that the most of students were very excited to learning corpus-assisted discourse analysis in classroom and mostly realized that there are much advantageous from learning corpus taught by the teacher. They seemed

very joyful and friendly corroborated with giving a positive attitude during the learning such as said by 1st participant “*it is much enthusiasm to have corpus learning because it has much benefits for searching within a whole text*”, in line with 2nd participant added that “*in my opinion, learning corpus is very simple and I am so happy to know that software because I can recognize word, phrase directly without reading word by word*”. Those particular responses emphasized that corpus tool is very useful for students in learning and researching any kinds of linguistic and most of responses realized that they could highlight the main content or the intention of the text as a presupposition, as proven by 3rd participant that she argued “*since I applied corpus tool by seeing concordance line, I can perceive the content directly instead of a long reading time*”, and 4th participant expressed “*I think I can catch the point of the message when I used corpus to know at least the meaning or intention of the writer*”.

Based on a representative sample of responses, the students as participants were able to operate the concordance software and they also could adapt with the theory such as they were able to search the word that assisted to find out microstructure element in this case. For instance, they applied searching for the first element of microstructure and so on in the following below: *the first, transitivity system*, the student understood the material and related to *functional grammar* course they had learnt and were able to applied the transitivity system by searching the word “*consider*, claim* argue*, state**, including their inflected form *-s, -es, -ed, etc.*” as verbal process through concordance software. Further, other processes such as behavioral, mental, causative, and relational processes, also implemented in the classroom in order to get a comprehensive understanding. Furthermore, the students were able to explore and analyze their frequencies, collocations, and concordance line from the

searched word. In the last of element analysis, eventually, most of students could interpret and create a first presupposition toward the writer’s purpose, intention, or ideology in making a text.

Secondly, **nominalization**, in this part, the most of students could optimize the corpus to show the nominalization where they had also insightful material due to have learnt a structure course in a previous term. They were able to modify the words that wanted to explore by adding suffixation *-ment, -ion, -ive, -or, -er, etc.* in each word as their derivational form. Moreover, they could recognize whether the actor is exist or not and they could know representatively a hidden actor in social practice among the discourse of text in the corpora. The attached caption below presents the corpus works by searching the word *government* with frequency, collocation and concordance line in the whole text.

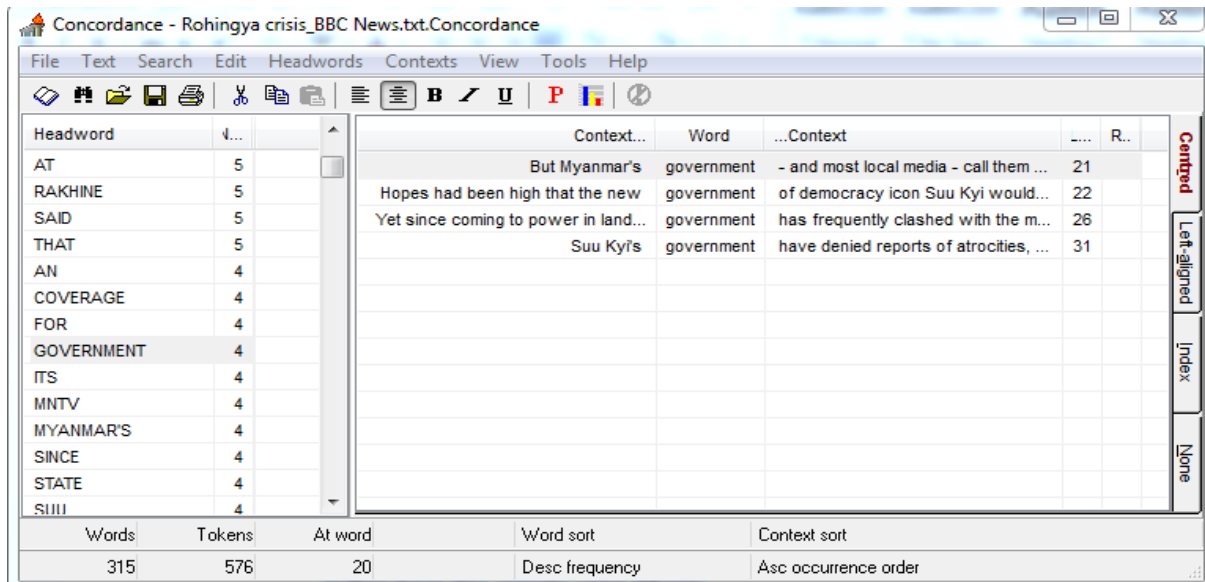


Figure 1. The word “government” occurrence, frequency, and concordance line as one of microstructure analysis

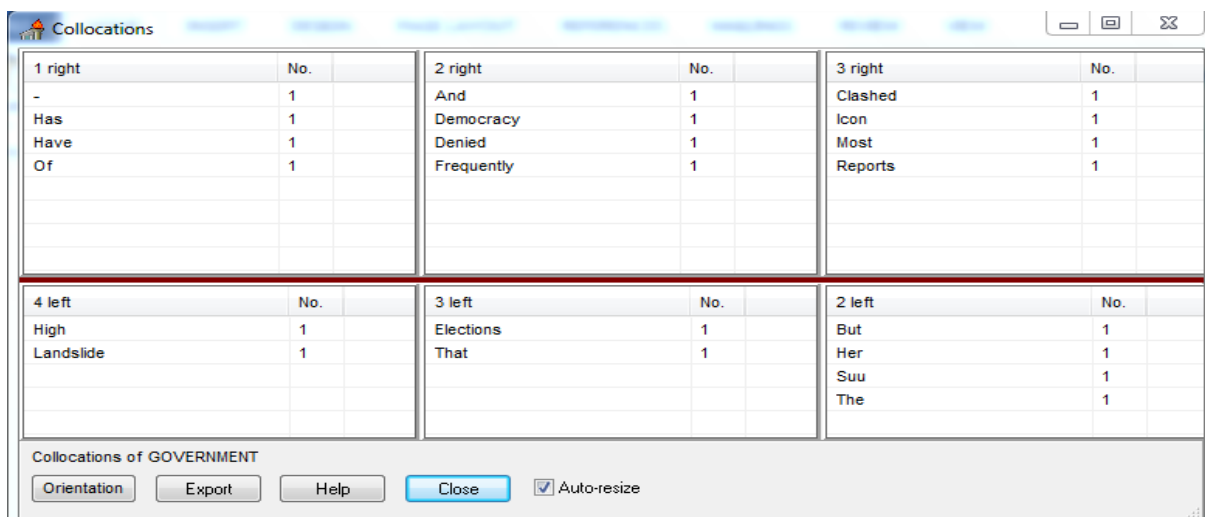


Figure 2. The word “government” with its collocation as one of microstructure analysis

From the mentioned figures above, it showed that the corpus can produce specifically the word ‘government’ and inform the students that *government* is seemingly general. As contextually which is related to that word, it cannot show and identify the real actor in that case. By adding suffix in the end of the word, it is hardly to investigate the identity of the writer or actor’s action. It is highly

tendentious to be not cleared in the public by the publisher of media. Thus, the corpus only can support in making the researcher’s claim interfacing with other aspects of sociocognitive by Van Dijk (2008).

The third is *passivation* where seemingly likes a same with nominalization due to this part only needs to involve the affixation such as suffix in the identification. The researcher explained to

the students that passivation always appears in the predicator or process. It is usually attached in the predicator by suffixing *-ed*. As semantically, the attachment can derive the meaning for instance *STOPPED* is the one of passivation that occurred in the corpora and indicated a hidden actor or subject due to the needed of passivation is only enough with the object not subject to catch the meaningful of the communication goal, but nevertheless, it always creates implicit meaning and had an ideology beyond the text.

Fourthly, the last element of microstructure analysis is *reference*. In this case, the researcher explicated the main of reference to the student in the classroom. It was aimed to show the strong relevance or correlation with textual meaning and mode system in SFL (Gerot & Wignell: 1994). The way to analyze the reference was finding out the name and criticize its associative within the corpora. The researcher demonstrated by separating to element of participants, namely; human and non-human. For the human category, the researcher just typed the names related to the case in KWIC such as OBAMA, TRUMP, etc. for the counterpart, it just needs to type either the name of party, place, or profession. Contextually, the word OBAMA or TRUMP have highly

associated with AMERICA and frequently collocated with the name of party, profession, and many names referred to them. From the result, the students were able to optimize the concordance line due to that feature can recognize the actor stance and its ideology in a particular corpora. By analyzing the reference element, it can strengthen the interpretation to the particular social practice adopted by sociocognitive (Van Dijk: 2009).

Students' responses on the implementation of corpus software in teaching CDA

Furthermore, the continuous part is to answer second aim of the study. It is delineating the students' responses toward the implementation of corpus software in teaching CDA and the divergent of leaning CDA by using corpus and conventional way as representatively. In addition to obtain students' responses, the questionnaires were distributed to 20 participants and consisted of two parts. Part one released about the teaching CDA covered five questions as part two mainly concerned to the implementation of corpus software in learning classroom covered five questions too and each part is summarized in the table below following Likert Scale.

Table 1: The summary of the questionnaires on teaching CDA through corpus based analysis

Statements	Category				
	A	B	C	D	E
1. Students' understanding on CDA material in classroom	11	5	3	1	-
2. Students' understanding on microstructure framework of sociocognitive	9	5	6	-	-
3. Students' work on microstructure analysis by grouping in conventional way	10	6	2	1	1
4. Students' interpretation on the result of microstructure analysis	4	12	2	2	-
5. Lecturers' evaluation on students' analysis	8	5	5	2	-
6. Incorporating corpus based approach with CDA material and demonstrating	9	7	3	1	-
7. Students' training on corpus software in classroom	9	4	4	2	1
8. Students' training on corpus-assisted CDA by grouping	13	6	-	1	-
9. Students interpretation on the result of microstructure from corpus software	2	10	6	2	-
10. Lecturer and students' evaluation on their work for the pedagogical implication	15	4	1	-	-
Total Score	90	64	32	12	2
Average Score (%)	45%	32%	16%	6%	1%

A = Very Good, B = Good, C = Enough, D = Bad, E = Very Bad

From the table 1 above, the part one is covered by question number 1-5. Essentially, the first statement emphasized

about the students' understanding to CDA material in general. As the result, it had good response from 11 students with the

score A categorized 'very good' whereas one student categorizes 'bad'. The second statement was students' understanding on sociocognitive focusing on microstructure model couched by Van Dijk (2008) and the third one was students' analysis on microstructure analysis by grouping in conventional way that corroborated with both of second and third statements had 'good' scores although the some of students felt easier to understand the theory than the analysis. The fourth statement was talking about the students' interpretation on what they had analyzed while the fifth one was about lecturer's evaluation. In accordance with the result, the fourth score was better than fifth one with a 'good' one, but nevertheless, four student gave 'bad' category. The overall of evaluation on the teaching CDA in classroom showed that most of students were satisfied and excited with joyful that proven by the score 'very good' categories from the first-fifth questions.

Moreover, the sixth statement to tenth one was part two of further category. It delineated the implementation of corpus tool in classroom to the students. In accordance with aforementioned table 2, the sixth statement aimed to incorporate the corpus tool, named concordance software, with CDA material in analyzing Van Dijk (2008) framework and it was about demonstration of corpus interfaced with CDA material. Half of students' responses toward this case were 'very good' categories. It meant that half of students were understood and interested in learning CDA through corpus tool. Further, the seventh was about students' training on corpus tool individually with searching any words, phrases while exploring their frequencies, collocation, and concordance lines. In that case, students gave the score of 'good' category too. In line with previous case, the eighth statement provided training session that the student required to train their analysis of microstructure through corpora. Most of

students realized that they felt easier with a 'very good' score from 13 students, 'good' one categorized by 6 students but nevertheless, there was one student who gave 'bad' category. The continuous statement was interpretation session that was aimed to train students more criticism regarding the result of corpus analysis and it was categorized 'good' by students due to they felt corpus tool caused them easier in interpreting the corpora of a specialized issue. The last statement was lecturer and students' evaluation. On the other hand, it was intended to discuss students' work together for the evaluation and implicational pedagogy such as the present study is an essential method to teaching, learning, and researching to be an innovative way for teacher, students, and researcher. Therefore, the category of tenth statement was totally good and successful corroborating to the attached table above. This was strengthened by the overall result from the first-tenth statements that 'very good' category had total score 90 or average score was 45%, 'good' category has 64 score or its average one is 32%, 'enough' category had 32 score or 16% for average one, 'bad' category was 12 or 6% average score, and it was totally different for 'very bad' category that only had 2 or it was only 1% for that category.

CONCLUSION

This study has discussed the exploration of teaching CDA through corpus tool in classroom. it can be noted that teaching CDA through corpus tool can be make more interesting, effective, and easier to research the words, phrases, contexts by their occurrence, collocations, and concordance line analysis in contributing to reveal some cases of discourse for society perspective. Accordingly, this present study also can emphasize that CDA can be adapted with teaching aspect and it is possible that other linguistics studies can be applied too into pedagogical aspect. Some of results corroborate and reinforce the findings of

previous studies on incorporation of CDA with corpus studies that this present study is supported and had been strengthening each other as stated by Charles, Peccrari, & Hunston (2009). It is strongly drawn that it is somehow not only can be used for general discourse analysis but also for specific CDA such as relates to investigating three levels of analysis from Van Dijk (2008) framework.

Further, the implicational pedagogy based on findings of this study can be delineated. *First*, learning corpora can make students be aware to a real language production and authentic language of social use. *Second*, students can optimize the corpora for any kinds of studies and interface with other studies due to corpora definitely enables to make it general corpora or specialized corpora based on the need. Thus, the students can utilize the corpus software by free charge or offline way. *Third*, it is emphasizing that corpus linguistics can be used for language evaluation for pedagogical aspects such as recontextualizing English textbook, language curriculum, students' task feedback, or any designing of language learning in classroom. however, corpus software must be supported with other studies to reach the purpose of the research due to it just highlights a surface analysis as a first step of further analysis.

REFERENCES

- Adel, A., & Reppen, R. (2008). *Corpora and discourse: The challenges of different settings*. Amsterdam: Benjamins.
- Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh: Edinburgh University Press.
- Biber, D. (1998). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- (2006). *University Language*. Amsterdam: Benjamins.
- Charles, M., Diane Peccrari, & Susan Hunston. (2009). *Academic writing: at the interface of corpus and discourse*. New York: Continuum International Publishing Group.
- Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics*, 22, 75-95.
- Gerot, L. & Peter Wignell. (1995). *Making sense of functional grammar*. Queensland: AEE Publishing.
- Hou, Yu. (2014). *A corpus-based study of nominalization in translations of Chinese literary prose*. Bern: Peter Lang.
- John, T. 1991. From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. *ELR Journal*, 4, 27-45.
- McEnery, T. & Wilson, A. (1996). *Corpus linguistics*. Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R. & Cheepen, C. (2006). *Corpus-based language studies: An advanced recourses book*. London: Routledge.
- Swales, J.M. (2002). Integrated and fragmented worlds: EAP materials and corpus linguistics. In J. Flowerder (Ed.), *Academic Discourse* (pp. 150-164). London: Longman.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Van Dijk, T.A. (2008). *Discourse and context: A Sociocognitive approach*. Cambridge: Cambridge University Press.
- Vyatkina, Nina. (2016). Data-driven learning of collocation: Learner performance, proficiency, and perceptions. *Language Learning and Technology Journal*, Vol 20, 3, pp. 159-179.
- Wodak, Ruth & Michael Meyer. (2009). *Methods of critical discourse analysis: Second edition*. London: SAGE Publication Ltd.