Teaching the Construction of Feminist Ideology to EFL Learners: A Corpus Approach

Mamona Yasmin Khan*  Nausheen Rasheed †  Shaheen Rasheed ‡

Abstract

Corpus Linguistics has gained much importance and is now continuously been applied in linguistics as well as, in literature. This paper aims to investigate the use of corpus linguistics in teaching literature to students. This research has used both qualitative and quantitative approach. The question it probes to find out is, how Simple Concordance Program (Smp 4.09) aids in teaching feminist ideology to EFL learners? For this study, the data have been selected from Leyshon’s novel The Color of Milk 2012 through purposive sampling technique. The results of this research illustrate: 6 different types of words (tokens), which were supporting the feminist ideology, have been extracted from the word list of the sample and then, feminist analysis has been carried out, according to their occurrence in the phrase and sentences. For future researches, it is suggested to conduct a stylistic of the same novel by the application of corpus tool.

Key Words: Corpus Linguistics, Word list, Key Words, Concordance, Simple Concordance Program (4.09)

Introduction

In this modern age, the education is being computerized. Technologies are being introduced in the educational set up in order to facilitate teachers and student and make smart learning possible. The traditional methods of teaching English language have now altered into modern one which is directing the teachers towards the importance of corpus linguistics. Corpus linguistics is the study of the samples of the naturally occurring material i.e., language. These samples, when stored in computers electronically, become digitalized and are called “corpus”. According to Oxford Dictionary, “Corpus is a collection of written or spoken texts”. Biber (2011) states that Corpus Linguistics is that approach which is concerned with the description of language use. In other words, this study is focused on how the language is used in everyday conversation. It can be written such as newspaper, pamphlets, leaflets, books, magazines etc. and it can be a spoken discourse, which is transcribed first and then, electronically stored. Corpus linguistics is needed in order to study a language in depth and stylistically. Hence, it is beneficial to do linguistic analysis. Moreover, corpus linguistics has also its uses in literature. Different researches have been done on the application of corpus linguistics in literature which is also termed as Corpus Stylistics (Biber, 2011). Literature can be taught to the students, by using the corpus tool i.e. Simple Concordance program.

In our previous research paper, “Teaching English through Literature to EFL Learners: A Corpus Approach”, we have raised the point of using corpus tool in English language teaching in EFL classroom. We have highlighted the fact that the

*Assistant Professor, Department of English, The Women University Multan, Punjab, Pakistan.
Email: mamona6231@wum.edu.pk
† Scholar, Department of English, The Women University Multan, Punjab, Pakistan.
‡ Scholar, Department of English, Bahauddin Zakariya University Multan, Punjab, Pakistan.
corpus tool can help the teacher in teaching and assisting students in learning English language in a comfortable environment. The English language which is being taught in our country is through a dull way which is why students find it a rough, tough and strenuous job to learn English language in dry environment. The reason is that the way the grammar is being taught to the students is not innovative enough to grab their attention and make them love the process of learning. Instead, at every next step, the teacher can feel their students holding their breath in fear and forgetting about letting it out, as the teacher goes on explaining the grammatical structures. This is why, we found it imperative and the hour of need to make the teaching of grammar somewhat interesting. In Khan, Rasheed and Rasheed’s (2021) paper, it has been suggested to employ Scp (4.09) tool. The benefits of using this tool are: Firstly, it is free of cost window-based program which can be downloaded and transferred to the other computer devices. The teachers and student can easily access the program without paying a single penny which is needed for the subscription of other corpus tools. Secondly, as soon as it is downloaded in the computer/laptop it remains there. There is no time limitation allotted for its usage. Thirdly, the program is easy to handle for the beginners. The students do not have to be bothered about the ways of operating the corpus tool. It is quite easy to understand and learn the features of Scp (4.09) and operate any kind of written text.

In general, all the corpus tools can read the written as well as spoken text which is transcribed and added in the system to be read. The corpus tool makes the wordlist of the words which have been used in the text. The wordlist presents frequency (the number of times a particular word has appeared in the text). Moreover, by selecting a word e.g., ‘for’ and putting it in a search engine, the location or the occurrence of the selected word within in the text can be traced. Hunston’s (2002) following comments increases the credibility of the corpus which are: firstly, the corpus can give the frequency of the words used in a text whether the verbs are present in written text or in spoken (transcribed) text. All are presented on the screen with the help of a corpus tool. Secondly, the words which occur together i.e., collocative words are also identified with the help of corpus tool. Thirdly, semantic prosody can also be determined by corpus tool and fourthly, the insights relevant to the spoken and written text that how they are distinguished with respect to the context are also attained. Hunston and Francis (2000) add that the corpus tool in short can perform five tasks which are, “a corpus tool selects, sorts, matches, counts and calculates”.

The corpus approach has been taken up in this research due to two legitimate reasons which are somehow interdependent. The first one is that the text which has been taken for analysis in this study has been taken out from Nell Leyshon’s novel The Color of Milk (2012). The novel is feminist in nature and demand feminist analysis. In feminist writing, a repetition of certain words and a specific repertoire is experienced. The words which have different meaning in other context can have different meaning in the feminist context. Due to certain gender specific labels, collocation, repetition of some words over and over again, increases the intensity and meaning of the word. For that purpose, the corpus approach has been preferred, as it has been aforementioned that the corpus tool helps in providing the entire list of the words and other characters used in the text in the form of a wordlist. It provides ease to the researcher and the teacher to pint out the location of the occurrence of the gender specific labels along with their frequencies, as it is not possible to manually count the frequencies of those feminist words. Hence, the wordlist eventually helps in highlighting the words which contain certain labels, derogative and belittling words which is the particularity of feminist writings. This is the reason of selecting corpus tool, in order to conduct a feminist analysis.

There are other numerous tools such as British National Corpus (BNC), WordSmith Tool and Corpus of Contemporary American English (COCA) available which can offer more features but those tools are not easily accessible. Those corpus tools require uninterrupted facility of internet. Secondly, those tools have their own stored corpora consisting of millions of words comprising of all sorts of words (written, spoken, speech, debate, newspapers and car manuals etc.). With the stored corpora the text can be compared, hence, a comparative study can be carried out. It can be deduced from the above discussion that if a researcher is willing to comparatively analyze his
text with the existing corpora of corpus tool, then, he must go for these tools (BNC, WordSmith, COCA). As the study of Romer suggests that the selected text is compared with a pre-existing data which he calls ‘software package’ while using WordSmith tool. After comparing it with the ‘software package’, the text is observed stylistically and learner finds out the linguistic competence of the writer, which is why it is also called Corpus Stylistics. Since, the Scp (4.09) does not contain a stored corpora so it does not compare the text with anything. It simply runs the text file, reads the text and provides the wordlist along with the frequency of the words used in the text. Neither, the Scp (4.09) offers a comparative analysis nor, it was the aim of this research. However, a comparative analysis of the texts will only be possible when the literary text belonging to same age but written by different writers is selected. This research only aimed to find out the list of words which have made the contribution in the formation of feminist text.

There are certain concepts and terms, which are used while doing corpus analysis, need to be defined and these are:

**Simple Concordance Program 4.09 (Scp 4.09)**
This program contains in-built alphabets relevant to many languages such as English, German, Polish, French etc. It can also generate the alphabet with the assistance of alphabet editor.

**Features of Scp**
Firstly, the words and phrases can be explained in concordance program. The KWIC (keywords in context) and line-based concordance are also achieved. Secondly, the files which are shown on the screen are saved in html file in the system. Thirdly, there is an option of arranging wordlists in ascending and descending order. There is also a feature of prefix and suffix in this program. Furthermore, the feature of stop list is also there which can help in removing a word from the list. There is also a characteristic of getting a statistical analysis of the text (Kuzyakov, 2020).

**Data Driven Learning (DDL)**
Tim Johns was the first who introduced DDL in his work. He opines that DDL provides assistance both the teacher and learners in language teaching and learning process (Hunston, 2002). Römer defines DDL (Data Driven Learning) as that type of learning method where students explore the language pattern on their own and find out the uniformities and peculiarities (if any) in those lexical patterns. Hunston (2002) describes the aims of DDL which are approaching the meaning of the text by extracting it from its context. In addition to it, DDL makes students use corpus data such as concordance lines or sentences by creating such situation.

**Summary of the Novel**
The script which is under analysis in this study has been extracted from the novel of Nell Leyshon “The Color of Milk”. The story is about a 15-year-old young girl, Mary who was used to live with her poor family at the farm. After some period, she was sent to the house of a pastor where she provided her services as a housekeeper. There she looked after the sick wife of the pastor and managed the kitchen. She became the live-in maid. Due to the severe illness of the mistress, she did not make it and died. There was another maid working it that house but she too left
the house. Now, there lived only Mary and the pastor. The man had his son Ralph who used to come over and threatened and tried to suppress the maid. The master, used to give a lesson to Mary on how to read and write. One day, while he was teaching her, he attacked her and then it became his pattern to come over every night in her room to rape her. Mary remained silent and did nothing. When things went out of control, she slit the throat of her master with the help of a cheesing cutting knife. After that, she ran to her farm back to be hid under the hay. The police found her guilty and then she was sentenced to be hanged till death. This is the novel actually written by her who at that moment was serving her time in jail and had mastered writing. As this novel is presented to be written by a young uneducated girl who was unfamiliar with the rules of English Grammar i.e., Structure construction, rule of capitalization, punctuation. Moreover, she created the words on her own such as, “somatt” from something. The novel has been observed to be a product of feminist writing that has been written by a person who was a disabled woman, a maid and a rape victim who was treated as an object of leisure time and were involved in having sexual relations out of her consent. The last important point to mention is that she was executed while she was bearing a baby in her belly as a result of that rape. All of these instances and fact have made this novel perfect for exploring the formation of feminist ideology in the text.

Statement of the Problem
A great number of studies have been done in which the task of English language teaching has been performed with the assistance of literature on which the corpus tool was applied. It has been observed that there is no work done on the application of corpus tool including Simple Concordance Program in order to teach a literary ideology to EFL learners.

Significance of the Study
This study is focused on the application of corpus tool for teaching the construction of feminist ideology to EFL learners. Previously, the researches which have been conducted were on the application of corpus tool on literature for teaching phrases, vocabulary or English Grammar with the help of corpus tools other than Scp (4.09). The novel which is under analysis has not been analyzed before neither from the feminist perspective nor with the help of any corpus analysis tool. Therefore, this research aims to introduce the use of Simple Concordance Program for teaching literary ideology in EFL classroom.

Research Question
In what ways Simple Concordance Program (Scp 4.09) aids in teaching feminist ideology to EFL learners?

Literature Review
Jones and Waller (2015) write that John Sinclair developed COBUILD in 1980 which was the first learner’s dictionary. This dictionary utilized and supported the use of corpus data because it can provide word frequencies, collocation and colligation.

Chambers and Sullivan (2004) call corpus approach as “electronic literacy”. They believe that it has a significant function to perform in a language teaching and learning process. The application of a corpus too has been witnessed in their study for teaching the writing skills in French language to their students. The findings reveal the efficacy of the use of corpus in getting feedbacks.

Meunier (2002) is in the favor of using corpus in EFL classroom. In her research, she has talked about the use of concordance and teaching forms. Keck (2012) also emphasizes the usefulness of corpus linguistics for the purpose of teaching second and foreign language to the students.

Khan, Rasheed and Rasheed’s (2021) study have applied the same tool Scp 4.09 on the poems, in order to teach grammatical structures of present indefinite, present continuous, past indefinite, imperative and conditional sentence structures. The findings reveal that the tool helped in providing the wordlist of verbs which were cooperating in the formation of aforementioned structures.

Römer in his paper (2006) has explained the use of corpus tool (WordSmith) in teaching language pattern as well as literary studies to the learners. He states that there is a connection and interconnectedness among literature, linguistics, pedagogy and corpus. As corpus tool can be applied on a literary text which is then linguistically assessed.
and then taught to the students. Hence, proving the connection amongst them.

Al-Rawi (2017) has applied corpus linguistics, in order to analyze Charles Dicken’s novel “A Tale of Two Cities”, in his article. His study has found out the stylistic use of keywords in Dicken’s novel, by using Laurence Anthony’s Antconc. He has applied both qualitative and quantitative approach in his research and compared the keywords of Victorian novel (the node) with 35 Victorian novels (as a reference corpus). His findings are not the generalized ones, because the number of keywords he used to analyze are not enough but, still he has provided a useful insight to the researchers. Mahlberg (2010) has also done qualitative and quantitative study of the nineteenth century fiction and applied corpus linguistics. In her paper, in order to know the style of writing of the author, she has focused on the keywords, suspension, annotations and clusters. In another research of Mahlberg and McIntyre (2011) both qualitative and quantitative approaches have been applied to Fleming’s novel “Casino Royale”. In their paper, they have found out the keywords, key semantic domains and the group of words which have helped in the formation of a fictional world e.g., characters, setting and thematic signals. Two categories have been identified in this research which are termed as text-centered and reader-centered. The results are, the words which do not follow the semantic domain, can be considered as reader-centered keywords, which assist in fulfilling specific textual functions. In El-Nabi, Ezzat and Abdeen’s (2018) corpus study “A Corpus-based Lexical Study of Contemporary Feminist Short-Story Writers: with Special Reference to Alice Munro” In this paper, they have tried to explore the lexical items which are used by the contemporary female Canadian short story writer Alice Munro that what lexical choices and items she selected, in order to show the crumbled and decayed condition of a woman in her novel. They have done the analysis in the terms of colocation, concordance and lexical sets by using AntConc 3.4.4. The research of Wasterlain and Gilquin (2020) is about the use of corpus tool (SiBol) which is specifically designed to conduct a corpus analysis. This tool contains about 1.5 million articles taken from 14 English newspapers, published during the period of 20 years from 1993 to 2013. The researcher has analyzed the use of feminist word and the synonymous words in terms of collocation and their presentation with negative connotation by media. The results illustrate that during that period, the media presented the feminist movement in a mortifying manner. For over a long period of time, the negative connotation has been seemed to be attached with feminism. The analysis used tabloid and broadsheet which helped in providing objectivity to the research. The researcher also applied CDA to find out the framing technique employed by the media to misrepresent the term feminist and other synonymous words. The researcher concludes his research on this note that feminism is not what it has been presented over these years. It is a word choice of media to exploit it. A better representation is required in order to support it. Likewise, Lin (2021) applied WordSmith Tool to find out the characterization and carry out the thematic analysis of females presented in Drabble’s feminist novel “A Summer Bird-Cage”. The corpus analysis of the novel dictates that negative tone is provided to females when they share their perspectives about their lives with the readers.

Saricoban and Metin (2000) have shown the use of poems and folk songs in order to teach language to EFL learners. They are of the view that poems and folk songs are the entertaining genre which keep the students connected and satisfied with the learning process. Other researches such as Finch’s (2003) and Hussein’s (2004) have also suggested the used of literature (poem) for teaching the language grammar in EFL classroom.

Methodology
To do a corpus analysis is a time taking task. One cannot handle the data present in a great number. To analyze the whole Nell Leyshon’s novel “The Color of Milk”, would be a lengthy process so, the first two chapters were selected for analysis through purposive sampling technique. After that the selected text passed through the procedure of line break. In order to do the line breaking of the text, Shift+Enter at the end of each line were pressed. After that, the file was saved in a plain text form rather than saving it in a docs file. The reason of doing that is Simple Concordance Program 4.09 can only run the file which is in a “plain text” format and that is format of
Notepad. This action was followed by running the plain text file in the program which as a result created a wordlist, showing 1615-word vocabulary. In the next step, those words were selected which were promoting the feminist ideology and looked for their use in the phrase and sentences. Every selected word was manually entered in the keys of concordance. In the end, the feminist analysis of 6 selected words and their sentences in which they appeared, was carried out with the support of Tyson’s book, *A Critical Literary Theory* (2014)

**Analysis**

**‘Baby’**

![Figure 1: The concordance of ‘Baby’](image)

1) *when i was a baby they tied it to some piece of wood to straighten it*
2) *me and said the person what tastes a new cheese first will get a baby. / so i put it in my mouth*
3) *i waited till i swallowed then asked, how will i get a baby from that /
4) *mother, i asked, if i had a baby, would it have a leg like mine?*
5) *she held up the last one what was the size of a baby. and then she laid /
6) *i had a baby, she said, only i was on my own when he was born and there*

The appearance of the word baby in the first two lines indicates, the traditional gender role. The roles which are socially constructed by the society for the genders i.e., male and female. According to Lois Tyson, she says in her book “A Critical Literary Theory” that there are certain roles which are constructed by the society and these are, man is supposed to be tough, powerful, dominant etc. while, a woman is supposed to be weak, fragile, passive. A man is supposed to do a job and earn the money and woman is expected to marry a man and bear him children. To produce a baby is the most important task for a woman. So, the first two lines clearly show this importance. A girl, who is fifteen years old, who has no other knowledge of world, is instilled with the only knowledge by her parents and that is a person who tastes the cheese first, will get a baby. Again, in the last two lines, there is a traditional gender role. The cheese maker is a woman and who is slicing and distributing it is also a woman. The next two lines, show the identity crisis from which Mary is passing.
She was born with a twisted leg. From that on, she has been stigmatized. Now, she thinks the baby she will get, would that baby have the same twisted leg like hers. This is a piece of conversation which is going on between Mary and her mother. The writer has used baby six times in this conversation, also projects the value of having a baby for a woman.

'Sheep'

1) though sometimes the small sheep would come in if they lost their
2) that same day the sheep needed moving up to the field by the church what
3) hope was the one what went ahead and shut the gates so the sheep didn’t
4) send them through. and the sheep went in. and we closed the
5) we brought the sheep up, hope said. /i guessed
6) him i’m sorting one of the sheep. /he won’t believe me,
7) /she stayed to sort out a sheep, i said. /and then even
8) than when we take the /sheep up to the top field by the

Normally, the word “sheep” is used here is in a normal sense. But, in the line number (1, 6 and 7), sheep is used as a symbol for Ralph. Here, the word sheep is indicating a famous idiom “Black sheep in the flock” which implies that a least reputable member of the group and the one who brings disgrace and becomes the cause of insult for a family. Likewise, in these lines, Ralph is playing the role of a black sheep. So, sheep is used for foreshadowing the incident, that is going to happen later in the novel. In line # 1 Mary says, sometimes at night, some astray sheep comes in and they have to feed it. It means that women are sex object. They have to give man a sexual pleasure whenever they want and their consent does not matter. It is not like they have a choice. Similarly, in line # 6, when Hope says that tell father that she is sorting one of the sheep; it can be inferred that she completing her services that are imposed on her by that sheep.

Figure 2: The concordance of 'Sheep'


‘Tongue’

1) and i do know. but i can’t help it. cos i am how i am. my tongue goes
2) fast as the cat’s tongue when he laps up the milk from the bucket.
3) so i put it in my mouth and it was strong near enough to burn my tongue.
4) what is it? ain’t got nothin to say? devil got a hold of your tongue and
5) ralph laughed. you have a sharp tongue.
6) i got a normal tongue, i said. i stuck it out and he laughed again. you
7) if they have been sharpened upon your tongue they would be. i can see

At one instance, Ralph says to her that she has a very sharp tongue and later on, she is describing her tongue to Mrs. Graham that her tongue is very sharp. This reflects that a woman, is also inclined to accept what others say about her. The above extracted lines show that how Mary describes her tongue and how other i.e., Ralph talk about her tongue. Apparently, she seems to be not following the commands, which are given to her but, psychologically, she accepts that she has a sharp and witty tongue that moves very fast. This attribute is given to her by the society and she has admitted it as a fact.

‘Cheeks’

1) /bugger the jobs. get your cheeks back on the chair. / so
2) bed. he looked thin and his cheeks was / sinking in to his

Figure 3: The concordance of ‘Tongue’

Figure 4: The concordance of ‘Cheeks’
1) *bugger the jobs. get your cheeks back on the chair.*

2) *i sat down on the box by his bed. he looked thin and his cheeks was*

The first line is said by Mary’s grandfather. When she tries to get up and get back to work then, her grandfather says to her “get your cheeks back on the chair”. In this line, objectification is done. Instead of saying sit down, he has done fragmentation of Mary’s body. So, here the concept of Male gaze is involved, as proposed by Laura Mulvey. The fragmentation of female body is done by the males and these fragments are then observed as a sexual object. That is how objectification is done. This is their normal discourse. In the second line, the word “cheeks” is used normally, in the sense of the part of a face. So, it is a sexist discourse shown in these lines. The word “cheeks”, which is used by grandfather is different than the same word which is used by Mary for him.

‘Lump’

The word “lump” is used in three different ways. In line # 2, it means, the adam’s apple which is present in some people’s neck and in line # 3, lump is used in a sense of pile of cheese which Mrs. Graham was eating. In the line # 1, Mary has used it, in order to refer it to herself. When Mrs. Graham asks her about her birth, she replies that she was like a lump in her mother’s belly, which made it difficult for her mother to bend and do work in the farm. She thinks of herself as an obstruction and a hurdle in the way of her mother. Here, the existential crisis of Mary is shown. She does not feel herself worthy and valuable. She thinks that she is a problem for her parents, especially for her mother, because she has a twisted leg and she cannot work the way normal healthy people do. Furthermore, she has equaled herself to an inanimate object. As the line # 3 shows, the lump is used to indicate the pile of cheese and Mary has called herself lump too thus, she has done dehumanization by herself.

**Figure 5**: The concordance of ‘Lump’
As the frequency shows, the letter i has appeared 1159 times in this sample. All the lines are difficult to be shown so, the screenshot of the first 33 lines have been taken. These lines illustrate that the writer of the story is an illiterate person. The reason behind is obvious. Before the waves of feminism started, women were not given many rights and a right to get proper education was amongst them. The only job and role that was assigned to them, was to stay at home, within the four walls and produce children as many as her husband wants. So, in the sample she has only used words in the small cases. She is unaware of the traditional grammar and writing style. The only thing that Mary wanted was to tell her story to the readers and she has done that even by breaking the rules of the grammar.

There is another reason of her using i in small case. It can be the identity crisis, from which Mary is suffering. She does not think of herself as an important being thus, she has used small i, to indicate the value of her existence in the world.

Moreover, if keenly observed, her character also seems to be breaking the gender stereotypes. She has broken all the norms, at first, from writing a novel, at that time when education was not for the girls. She
has written the novel, even being illiterate. She has also broken the traditional rules of grammar and wrote the novel in her own way.

**Conclusion**

This study illustrates that literary texts can be taught by using Corpus linguistics as a method. As the research shows, with the help of Scp (4.09) tool, 6 types of word which were promoting the feminist ideology, were extracted from the word list of the collected sample and analyzed on the basis of their use in the sentences. Hence, the utilization of Simple Concordance Program 4.09, has proved to be very helpful in teaching feminist ideology in EFL classroom.

**Future Recommendations**

For future researches, it is suggested to conduct a stylistic analysis of the same novel by using a corpus tool.
References


  https://www.theidioms.com/black-sheep/


  https://www.oxfordlearnersdictionaries.com/definition/english/corpus


  *The Encyclopedia of Applied Linguistics.*
  doi:10.1002/9781405198431.wbeal0256


  https://drp.su/en/catalog-soft/other/simple-concordance-program-113696

Lancaster.ac.uk. *Word Frequency Lists and Keyword Analysis.*
  https://www.lancaster.ac.uk/fss/courses/ling/corpus/blue/l03_top.htm


  https://wmtang.org/corpus-linguistics/corpus-linguistics/


Wasterlain, J., & Gilquin, G. “Feminism is a dirty word”: The media representation of feminism in a corpus of British newspapers (1993-2013).