

CORPUS LINGUISTICS FOR ESL MATERIAL DESIGNING: A QUALITATIVE STUDY BASED ON ANTCONC SOFTWARE

L. W. D. B. Pabasarani*

Institute of Technology University of Moratuwa, Sri Lanka

*bhagyap@itum.mrt.ac.lk

bhagyap@itum.mrt.ac.lk

ABSTRACT: Corpus Linguistics in the context of English as a Second Language (ESL) is high in demand as an approach that guides ESL practitioners in making informed decisions about language teaching. This discipline of Corpus Linguistics can be efficiently incorporated into material designing and teaching methodology, paving the way for both feasible and quality teaching and learning processes. The tool used for corpus analysis in this research is AntConc software. This research aims to unravel how the data derived from AntConc software can be used to make materials that reflect the real-life language use. The analysis derived from studying and analyzing a corpus is indeed an effective guide in the context of English Language Teaching (ELT). The text analysis performed by this software proves that the frequent patterns of language it identifies are a valuable resource for creating materials that accurately represent real-life language use.

Keywords: corpus linguistics, material designing

1. INTRODUCTION

According to Reppen (2010), a corpus is a 'large, principled collection of naturally occurring texts, (written or spoken) stored electronically'. This can be based on any actual written compilation of texts or transcribed spoken texts, collected for a particular purpose. The phrase 'naturally occurring texts' indicates the fact that the compilation must consist of authentic texts, which should indicate the real-life language use.

The next question raising worthwhile is what does it mean by Corpus Linguistics? Although it cannot be traced as a separate branch of Linguistics, it is proven to be beneficial in various other components of Linguistics, such as, Phonology, Phonetics, Morphology and Sociolinguistics. Corpus Linguistics is the research-based analysis of a given corpora which aims at decoding Corpus Linguistics properties such as Concordancing/Key Word in a Context (KWIC), Collocates, Multi-word expressions, Frequency lists, and Keyword lists. The analysis requires corpus analysis software such as AntConc or Compleat Lexical Tutor.

As Bennett (2010) points out there are two main research questions in Corpus Linguistics.

- What specific patterns are associated with lexical or grammatical features?
- How do these patterns can be differentiated within varieties and register?

Providing answers for the given questions can unravel the overarching as well as the underpinning connections between words, paving the way to understand the patterns of language.

Studies on Corpus Linguistics have long been held focusing on various concerns in the field of ESL. One such concern is about material designing and development and how the language prompted by materials in the classroom does not resonate with the language used for real life communication purposes. This is indicated by Johns (2022) stating that 'if compared to real data, textbooks do not always offer a realistic model of spoken language in terms of its forms and discourse feature.' The best way to bridge the gap between language used in the classroom and real-life language use is to use authentic materials for teaching. This is in fact one of the most crucially significant decisions that ESL practitioners should take. The aim of this paper is to incorporate the analysis of Corpus Linguistics with material designing, which promotes authentic or real-life language use. The analysis of the AntConc software can solve the matter if material designers can use the result of the analysis to develop or design materials which prompt the real-life language. The corpus analysis not only assists the ELT practitioners in making right choices in material designing but also provide chances

in cooperating authentic evidence in language teaching. In other words, the Corpus Linguistics analysis helps ELT practitioners to authenticate the language in a given context.

There are two main objectives in this research.

- To do an in-depth analysis of the feature; ‘Key Word in a Context’, (KWIC) by using the AntConc software specifically focusing on Conjunctions.
- To find different question types in which the analysis of the corpus can be utilized when designing materials for intermediate learners effectively.

There are many Corpus Linguistics tools such as Concordancing /KWIC, Collocates, Multi-word Expressions, Frequency Lists, and Keyword Lists. This research specifically focuses on Concordance of KWIC in order drive succinct results. In this approach, the targeted word class is Conjunctions. The focus is only targeted towards Conjunctions, since it assists in developing the materials based on Conjunctions in a more detailed manner. This choice is also made in order to maintain reliability.

2. METHODOLOGY

As the first step, seventeen articles from the magazine ‘Reader’s Digest’ were selected. The articles were selected aiming on a set of diverse concepts, thus it included current issues, entertainment, social issues, personal experience, inspirational stories and humour. This choice serves the fact of authentic as well as diverse data which must be included in a corpus. ‘Reader’s Digest’ magazine was a preferable option for this study since it is an internationally acclaimed publication. Next, the selected articles were converted as Ms. Word documents, creating a file for each article. Then, the meta data file was created by filling the details of the text number, genre code, text title, year of publication, publisher, Author’s name, gender, place of publication and the nationality. After that, the files were converted into plain text files. As the last step the Corpus Linguistic analysis was done. In order to do that, the selected corpus files were directed to the AntConc software. Even though the articles were converted into plain text files for clarity, the AntConc was able to read the Ms. Word documents in the KWIC option. Then, the analysis of KWIC was done by selecting specific words which were consisted of conjunctions and an intensive study was conducted searching the patterns of the language in right context and left context. The chosen conjunction was typed in the search query; thus, the results were shown indicating the word both in left context and right context. The target group of students aimed here are intermediate level students and to be specific, they are at upper intermediate level. According to the Common European Framework of Reference (CEFR) the students are at the level of B2. As the British Council depicts it, the ‘level B2 corresponds to independent users of the language, i.e., those who have the necessary fluency to communicate without effort with native speakers.

3. RESULTS

The selected Conjunctions for this research are ‘and’, ‘that’, ‘but’, ‘when’, and ‘because’. These Conjunctions were selected since they cater to the level of the learners who are intermediate and for the fact that they frequently appeared in the chosen texts.

The first key word analyzed in this research is ‘and’. The most frequently appeared word in the right context of the word ‘and’ appears as “and the” pattern. The next pattern which appears the most in the right context is the pattern of ‘and then’. As always it is used to indicate subsequent actions. At the same time the pattern of “and a + quantifier phrase” was quite prominent. The phrases noticeable were ‘and a pair of’, ‘and a whole lot of’ and ‘and a few’. The pattern of ‘and + Possessive Adjective’ such as ‘and my’, ‘and his’ could be deciphered. Also, the pattern of “and + Subject Pronoun” was visibly deciphered too.

The second keyword analysed in this research is ‘that’. When concerning the used word towards the right context, the main pattern which could be sorted out here was the pattern of ‘that the’. This pattern has appeared in contexts where the writer wants to showcase a distinctive idea. Just as with the keyword of ‘and’ the pattern of ‘that + possessive adjective’ and the pattern of ‘that + Subject Pronoun’ could be unravelled. At the same time, the specific patterns towards the left contexts were not designed according to specific words or phrases, thus it was difficult to decode any pattern.

The third key word discussed here is ‘but’. The most prominent pattern which could be decoded towards the right context was the pattern of ‘but + subject pronoun’, thus the patterns such as ‘but I’ and ‘but it’, appeared. At the same time the pattern of ‘but + possessive pronoun’ could be deciphered, provided with the example of ‘but my’. Just as with the case of ‘that’, the specific patterns towards the left contexts were not designed according to specific words or phrases, thus it was difficult to decode any pattern.

The fourth key word in this research is ‘when’. The most prominent pattern that could be deciphered in this word was ‘when + Subject Pronoun’. The most frequently seen phrase here was ‘when I’. At the same time in many cases the phrase ‘when + Subject pronoun’ was immediately followed by a verb. The next pattern which can be seen was ‘when + a+ noun’.

The last word analysed in this research is ‘because’. The first patten which could be deciphered in this pattern was the phrase ‘because of’. As always it indicates a reason for a particular condition. The pattern of ‘because + Subject Pronoun’ could be found here as well like in previous key words. The only verb which appears after because in this corpus was ‘because watching’ indicating the pattern of using the Present Participle of the verb in case of using a verb, after the conjunction; because.

4. DISCUSSION

As Reppen (2010) points out ‘knowing which words or which grammatical features are typical for a particular context of language use is a good starting point for instruction.’

Since the most common pattern in all five Conjunctions are ‘Conjunction + Subject Pronoun’ and ‘Conjunction + Possessive Adjective’, a gap filling activity can be given specifically focusing on the use of each structure under various contexts. A sample activity is given here.

Activity 1

Fill in the blanks, by selecting the correct phrase with the Conjunction.

1. She got very upset found out, she has failed her examination. (when she / that she)
2. She cries has failed the examination. (because she/ when she)
3. She promised will study well next time. (but she/ that she)
4. She studied well failed the examination. (but she/ and she)

The next significant word pattern is the use of Present Participle after the Conjunction of ‘because’. Here in this context, the Present Participle is not used as a verb, but as a Gerund. This prescriptive grammar rule can be easily taught when it is represented through common situations in the day-to-day context. For this purpose, a question-and-answer activity is taken since it stimulates the authentic language use. For this activity, the responses are expected as short and crisp utterances since in an informal spoken context, short and crisp utterances are normally used.

When considering how students get to learn the use of vocabulary, as Reppen (2010) point outs Schimitt’s (2004) notion about vocabulary, it indicates the importance of learning vocabulary ‘in

groups or chunks of language’, rather than learning them ‘in isolation’. The same concern can be applied here and the best part of it is that the learners not only learn about vocabulary or the grammar aspect of it but also about the context of its use.

Activity 2

Fill in the blanks by using the correct verb form to make meaningful sentences. Read out your response to the class.

1. A – Why are you silent?

B – Because with you can cause many problems.

Sample answers: (talking / chatting)

2. A – Why do you like to study?

B – Because Makes me happy.

Answer – studying

As Hunston (2002) points out, ‘corpora themselves can be exploited to produce language teaching materials, and can form the basis of new approaches to syllabus design and to methodology’.

The accuracy aspect of speaking can be promoted when the learners are instructed to speak using a few chosen structures in a controlled setting. Hence, the activity three is based on promoting the accuracy of speaking in a controlled setting. The students are expected to use the Conjunctions of, ‘and’, ‘that’, ‘but’, ‘when’, and ‘because’ in making accurate and meaningful sentences. The prompts to be used when speaking are given with the questions to assure an accurate outcome, which means that the students do not get a chance to be experimental with the language by taking risks. Since the chosen articles from ‘Reader’s Digest’ are appreciable sources for extensive reading, the students can be instructed to read an article or two before doing this activity. During the while reading phase the teacher can emphasize the Conjunctions. Then, the following activity can be done as a post-reading activity.

Activity 3

Join the two phrases given in the brackets under each sentence by using the Conjunctions of ‘and’, ‘that’, ‘but’, ‘when’, and ‘because’. Tell the correct sentence to the class.

1. What has happened to Mary? Can you tell me?

She had a fall, now she is crying.

Answer – She had a fall and now she is crying.

2. Why is she sad?

Her boss has told her she can no longer work with us.

Answer – Her boss has told her that she can no longer work with us.

3. How was the food yesterday?

We had a good spread many were spicy.

Answer – We had a good spread, but many were spicy.

4. When are you going to start working on the project?

Well, as usual it is just one day before the deadline.

Answer – Well. As usual when it is just one day before the deadline.

5. Why are you in a good mood?

Well, it is I am going to have a day-off tomorrow.

Answer – Well, it is because I am going to have a day-off tomorrow.

5. CONCLUSION

In conclusion, as corpus analysis helps in finding patterns common to a particular set of texts, analysing the corpora using a Corpus Linguistic software such as AntConc guides the ESL practitioners to take informed decisions about language teaching. The analysis can be utilized to design efficient materials. Besides, when the selected corpus is based on authentic materials, it becomes an added advantage since it effortlessly authenticates the language, while allowing the ESL practitioners to be more engaging and creative with material designing.

6. REFERENCES

- Benett.G. R. (2010). *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan ELT.
- Hunston. A S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press.
- Jones, C. (2022). *Authenticity in language teaching materials*. Routledge.
- Reppen, R. (2010). *Using Corpora in the Language Classroom*. Cambridge University Press