Use Of Corpus-Integrated Data-Driven Learning Approach In The Language Classroom: Malaysian Educator's Perspectives

By

*Hafizah Hajimia
Universiti Teknologi MARA, Malaysia
Email:hafizah.hajimia@uitm.edu.my

Manvender Kaur Sarjit Singh
Universiti Utara Malaysia, Malaysia

Nor Azira Mohd Radzi
Universiti Teknologi MARA, Malaysia

Nur Rasyidah Mohd Nordin
Universiti Utara Malaysia, Malaysia

Nur Asyrani Che Ismail
Universiti Teknologi MARA, Malaysia
Monther Mohammad Yousef Aljarrah

Abstract

Educators play a vital role because they are the key link between research and pedagogy. They are the most important element in increasing the development of corpora in language classrooms. Despite the development, not many educators seem to be using corpora in their classrooms, especially in Malaysia. Educator's perspective on teaching with corpora is an area that has mostly remained untouched. Thus, understanding the challenges and possibilities of corpora as seen by language teachers who aren't corpus experts is critical in moving the issue of corpora's use in language classes forward. The present study reports the advantages and disadvantages of CiDDL-PD from an educator's perspective. Seven (7) Malaysian University educators' perceptions of the potential implementation of DDL within the Malaysian L2 English teaching context were explored through open-ended interview data. Their understanding of the corpora was also tested through the practical application of the MyCoronas corpus. The findings indicate that the educators accepted the corpus approach in teaching and learning language in classrooms and their understanding of the new approach was reflected in the practical application. Based on the thematic analysis, six (6) advantages and two (2) disadvantages were identified. Qualitative analyses indicate that most of the respondents lack prior knowledge in applying corpus-based teaching aids. However, upon exposure, their awareness of incorporating CiDDL-Pd in creating corpus-based lessons increased in terms of their efficiency, knowledge and skills.

Introduction

Data-driven learning, or DDL for short, is an approach often used for second or foreign language learning. Commonly, language teaching and learning in the classroom are based solely on textbooks and the educators, whereas, in data-driven learning, both the educators and the students are treated as researchers who undertake guided instruction and task completion. Classroom concordancing or DDL is a corpus-based approach to language learning proposed by Johns (1991a). The Identify- Classify-Generalise technique was commonly used with this approach.

language-learner is also, essentially, a research worker whose learning needs to be

driven by access to linguistic data" (Johns, 1991a, p.2).

What he meant by the statement is that learners should be responsible for their learning; that is, "they should become learner-centred or autonomous learners through discovery learning" (Hajimia et al., 2019, 2022). On the other hand, educators play a vital role in creating a learner-centred environment because they are the key link between research and pedagogy. They are the most important element in increasing the development of various data-driven learning strategies including the development of various corpora in language classrooms.

The teaching and learning of English for specific purposes (ESP) is one of DDL's most significant contributions. The most well-known applications of corpus linguistics are in the study of language form and function with an emphasis on lexico-grammatical patterns or collocations, often known as the specification of language for certain purposes (Aston, 1998; Pearson, 1998; Sinclair, 2004a; Conrad, 2005; Belcher, 2006). Conrad (2005, p.399) stated that "teachers and students of a particular variety want to know the features of that variety, and so analysis of a corpus of that variety is clearly useful."

Concordances can also assist ESP instructors and students in understanding the vocabulary, ideas, and pragmatics of technical languages (Nolte et al., 2018; Yunus et al., 2016). ESP researchers and practitioners suggested using specialised ESP corpora to better match the needs of a small group in terms of building both declarative and procedural knowledge (Aston, 1996; Jabbour, 2001; Sinclair, 2004a; Gavioli, 2005;). Having a solid foundation in grammatical constructions and how they function in the text can greatly help language awareness. According to Gavioli (2005) and Hunston (2002), the topic of "what to teach" is far more significant and important for ESP practitioners than the question of "how to teach."

As Yunus (2017) indicated, in the present era of technology-enhanced teaching and learning, computer technology has revolutionised the fields of linguistic research (descriptive linguistics) and applied linguistics (language teaching and learning) with the advent of corpus linguistics" (p.137). For decades, Corpora have been utilised in language training. This learning technique has increased vastly because empirical research has regularly shown it to be effective. However, there is still a lack of use of corpora, notably among non-native English speakers, those with lower levels of proficiency, and those who are not in academic settings. This is ironic because DDL offers a lot of potentials "for developing modular flipped content, especially for online, hybrid, and remote courses" (Yunus, 2017, p.35).

Problem Statement

The traditional language teaching technique for teaching grammar has resulted in L2 learners' competency being deficient. The lack of necessary direction for effectively learning grammar explains why learners still lack an understanding of grammatical constituents and their application. Data-driven learning (DDL) has recently been found to help assist language teachers and learners in creating pedagogical approaches and language acquisition in research.

The goal of this study is to promote the use of corpus-integrated data-driven language learning pedagogy (CiDDL-Pd) in language classrooms. The researchers intend to investigate educators' perceptions of using the CiDDL-Pd approach in their classrooms. Using corpora to teach a language can also be called data-driven learning (DDL). John,1991 defines DDL as enabling language teachers and learners to have direct contact with "corpus citations and other information derived from corpora" (Leńko-Szymańska, 2014). "An important restriction of

education is that teachers cannot simply transmit knowledge to students, but students need to actively construct knowledge in their minds." (Bada & Olusegun, 2015) This method of learning a language is in accordance with the constructivism theory. Thus, this study aims to obtain educators' perspectives on corpus-based language analysis, evaluate the ability to compile corpora and using concordancing software, identify ways of using corpora in the language classroom and produce CiDDL teaching materials.

Research Objective

1.   To identify the advantages and disadvantages of Corpus Intergrated Data-Driven Learning (CiDDL)
2.   To identify the extend of difference between Malaysian scholars perspectives regarding CiDDL compared to other scholars.

Research Questions
1   What are the advantages of using corpora in L2?
2   What are the disadvantages of using corpora in L2?
3   Is there a significant difference between Malaysian educator's perspectives compared to other scholars?

Literature Review

A new paradigm of language education in the 21st century with DDL
        Language education in the 21st century provides limitless access to various teaching and learning methods, especially with technology incorporation like the data-driven learning (DDL) approach. The integration of DDL into language teaching and learning has made the education process more meaningful. This can be done through the application of corpora and concordance software (Gilquin & Granger, 2010; Boulton, 2017) which significantly contribute to more effective language teaching and learning as well as the development of several important skills. As mentioned by (O'sulliven, 2007, p. 277, as cited in Gilquin and Granger, 2010), "predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorising, hypothesising, and verifying" are among the skills which are used to explore language via the application of DDL and these skills can be further developed by the learners. Besides, language education nowadays has set a complex environment progressively for educators and students to engage in which higher-order thinking skills such as creative thinking, decision-making and problem-solving are highly valued abilities needed to succeed in it (Czerkawski & Berti, 2020). Therefore, these are achievable from the realisation of DDL in the teaching and learning process.

        Since this method uses authentic language materials to "empower both teachers and students to develop competencies in moving away from mere surface features of a text to selecting and understanding meanings and structures," Data-Driven Learning (DDL) appears to be the best solution for the development of metalinguistic knowledge and learner autonomy (Corino and Onesti 2019). Another key advantage of using corpora in L2 learning research is that they can provide empirical proof of authentic language use not only when compared to native speakers, but also when compared to the learners themselves. This is the case with learner corpora, where teachers and researchers may find the most common flaws and traits that distinguish native from non-native language usage (Granger 1996; 2015)cited in Forti and Spina (2019).

To be usable and valuable for second language learners, corpora must meet certain basic requirements. To begin, the corpus's texts must be appropriate for the learner's needs. These specifications are mostly concerned with the text's difficulty: reference Corpora are the most widely constructed and used corpora in any language, but they are only useful to advanced learners, who make up a small percentage of the language learner population because they are built to accurately describe a language from the perspective of native speakers. In contrast, the texts in a learner-friendly corpus would need to be acceptable in terms of competency level, reflecting either one broad level or various degrees of proficiency or difficulty. The texts in a learner-friendly corpus should ideally mirror the learners' interests, encouraging them to use the tool even more.

Direct And Indirect Applications

Without a question, the most evident, as well as the most important, role of a language corpus is to give knowledge about a language. As previously stated, the goal of the corpus determines what kind of texts are utilised to construct it and, as a result, what sorts of language forms can be predicted. The linguistic corpus notion ushered in a major shift in English language teaching and learning methods and philosophy (Binkai, 2012) cited in Jamal, Shafqat and Afzal (2021). A linguistic corpus is a collection of real-world language stored in an electronic database. It is offered as written or spoken messages on the internet or through software.

Teachers' jobs need them to connect with their students regularly. Starting with lectures and seminars and ending with visits with their parents, this official, yet unique, connection is built on teamwork. There are also tangible examples of this cooperation in the form of examinations, essays, or exercises that have a dual function: on the one hand, they demonstrate the student's level of knowledge and competencies. Besides that, they highlight mistakes and problems that need to be fixed.

Such proof can be found in a corpus, which simply compiles language data (without any personal detail). To go over the issues once again, other teaching tools might be made using this collection. Additionally, the frequency and size of particular errors may point to the need for further review and adjustment of teaching curricula or even software to better meet the demands of students in the actual world.

The topic of enhancing a teacher's abilities and energising the teaching and learning process is another one that is intimately related to corpus analysis. Some educators are capable of building their corpus. In reality, this means first becoming familiar with the software and programmes that can help with the production of such a corpus (improving tech skills, discovering new apps required to form a corpus), and then gathering texts that provide language data for the corpus (developing research skills).

Conversely, educators who don't want to construct their corpora might look for the information they need in already-existing sources (developing analytical skills). The most challenging task for educators is still to deliver specific instructions to their lectures and to support their students via data search and interpretation because language corpora just provide linguistic data without any analysis. As a consequence, the main responsibility of educators (and later pupils) is to evaluate the information gathered.

Applications of corpora in language teaching

Corpora are computer-readable, structured, and large archives or sets of texts that enable the examination of patterns of language use across multiple settings (Szudarski, 2018).

Since their inception in the 1960s, corpora have been recognised and used for language training, initially "in indirect and subsequently in direct uses" (Römer, 2011, p. 207). In the era of online, hybrid and remote teaching and learning, corpus-based teaching aids can supplement and enrich conventional curricula with virtual, flipped, and modular resources.

In CiDDL-Pd interventions, concordances and language patterns lexical and grammatical forms are commonly investigated by the learners. Teaching with corpora could be conducted both online and offline depending on the educator's creativity. By using the online hands-on tools provided, students could search a corpus database digitally either by themselves or with the help of the educators. In terms of the hands-off apps, the concordances and results of the patterns could be produced beforehand by the educators by printing or showing them in the slides. In the learning outcome, DDL was contrasted with more conventional, often deductive teaching strategies or various DDL techniques namely the hands-off and hands-on techniques. Improvements in learning have been evaluated using various forms of test questions and evaluations. Daskalovska (2015) found that "Macedonian university students who used concordances online learned English verb-adverb collocations more quickly than their peers who used deductive textbook resources (p.14) as well as its process via questionnaires, interviews, observations, eye tracking, and screen captures" (Daskalovska, 2015, p.15). For instance, Benavides (2015) managed to get feedback ranging from good to extremely negative, from the higher education students regarding their knowledge and experience with learning language patterns using their Spanish corpus.

Even though DDL has numerous scientifically validated benefits, it is underutilised by language teachers. Chambers (2019, p. 460) continued to emphasise the importance of closing the gap with the research practices. This gap has several causes. To begin with, over 90% of DDL research has focused on higher education students who converse with English as a foreign or second language and have various levels of language skills. This characteristic has hampered the distribution of DDL techniques to LOTE (Languages Other Than English) low-proficiency and early education learners. Furthermore, there is a lack of open availability of most representative corpora and corpus tools used in DDL studies. For instance, while having a large number of corpora and being frequently used as a corpus search and analysis tool, Sketch Engine "is only freely available to scholars at a select few academic institutions" (Kilgarriff et al., 2014, p.26). Third, the interface used to construct most corpora was designed by corpus linguists for corpus specialists, making it difficult for language instructors and students to utilise. "Many of the readily accessible corpora were produced as tools for linguistic research and not with instructional purposes in mind," (Braun, 2007, p. 308). Although not all DDL researchers are corpus linguists, they all spent a significant amount of time and effort examining corpora before planning and performing their educational experiments, "something that teachers who do not do DDL research are unlikely to do" (Hajimia et al., 2022).

For decades, unconsciously, language education has been embedded with corpora. Word frequency lists based on corpora have influenced the lessons and material development, and both educators and students have used identified examples in a language and its pattern by adapting corpora. Both in the education fraternity have gained from this inductive DDL applications, in terms of vocabulary acquisition and grammatical knowledge. They have also regularly outperformed non-DDL teaching approaches in terms of student autonomy and personalised learning proficiencies. These characteristics of DDL make it "ideal for use in hybrid, remote, and online courses" (Hajimia et al., 2019,2022). To augment and improvise existing pedagogy with digital and flipped subject materials, corpus-based modules can be created. By using their computers or other digital platforms, and at their speed, students can

perform corpus searches by themselves, and then submit their findings to their teachers using worksheets or other traditional media. This kind of activity could also help kids achieve wider educational goals including critical thinking, analytical abilities, and digital literacy (Vyatkina, 2020).

Methodology

Using a one-group quasi-experimental research design, the present study reports on the output of workshops and lesson planning activities conducted to enhance the use of CiDDL-Pd in language classrooms, among Malaysian university educators. Educators' perceptions of the training and the potential implementation of DDL within the Malaysian L2 English teaching context were explored through open-ended interview data. Their understanding of the corpora was also tested through practical application of the corpora. It is hoped that the educators would be able to integrate these materials and approaches into future teaching practices. This article reports a qualitative findings using the thematic analysis.

Research Method

Sampling and procedure

Respondents are seven lecturers from Universiti Teknologi MARA and Universiti Utara Malaysia. Open-ended interview questions were conducted in 3 stages. Stage 1 is before the training, stage 2 is during the training and stage 3 is after the training. This study will be showing the respondents' responses in the 3 stages. Finally, a few sample lesson plans were provided.

The data for this study were retrieved from open and closed-ended interview questions. The instrument which is the interview questions were adapted from Qiong, H. (2007) who conducted a study on the retrospect and prospect of Corpora and English teaching and Breyer, Y. A. (2011) who conducted a study on Potential, Evaluation, Challenges of Corpora in Language Teaching and Learning. These two were chosen because of their high number of citations and their suitability for this research.

The workshops were conducted by adapting Mukherjee's (2004) proposed framework for a five-day workshop. It consists of three main, sequential thematic modules, namely:

"Teaching about corpora through a brief presentation of the basic concepts, exploiting corpora to teach language by various hands-on explorations using coronas and corpus tools, and teaching to exploit corpora via discussion about classroom applications of corpora, supported by books relevant to the field of study" (Mukherjee, 2004, p.46).

Corpus tools

Corpus tools used for this study were the existing MyCORONAS (Malaysian Corpus of Online Newspaper Articles) (Hajimia, et al., 2021) and the computer-assisted corpus analysis (CACA) approach (Singh, M. K. S., 2010). Besides that, AntConc 4.0 by Laurence Anthony and Claws Tagger by Paul Rayson of Lancaster University were used as a tool to conduct the concordance search and part of speech tagging.

Data Collection and Analysis

Data was collected using the open-ended interview questions adapting Leńko-Szymańska (2014) questionnaire. The findings were analysed using a thematic analysis adapting Braun and Clarks (2006) stages of conducting a thematic analysis.

Findings

Question 1: Awareness about corpora and DDL

Before the training began, the educators' existing knowledge of corpus linguistics and data-driven learning were questioned. They were asked whether they are familiar with the term and concept of corpus linguistics and data-driven learning.

Based on the findings 5 respondents said that they were not familiar with it. This was followed by 1 respondent who seemed unsure. The respondents said, "not really familiar" (R3). Only 1 respondent had some idea about this term. He said, "yes, in a different perception" (R4).

R4 was further probed, "different perspectives… hmmm… I know its mainly for research" (R4).

This finding is on par with Chambers (2019) who mentioned that corpora are underutilized by teachers. From the Malaysian context, Yunus (2014), said that Malaysian educators are still not exposed to corpus linguistics or data-driven learning. This finding validates the gap identified in this study.

Q2: Advantages and disadvantages of Using CiDDL-Pd

From your experiences what are the benefits and disadvantages of teaching and learning using corpora?

2.1 Advantages
2.1.1a Theme: Easy to use, helpful with big data

Based on the findings, 4 respondents said that the coronas corpus was useful (R1, R2, R3, R6). R1 said that it is useful because it's very simple to create and yet there are free apps for this. R2 said it can help the teachers. Finally, R3 said that it is helpful especially when the data is enormous.

Based on the response words like 'simple to create', ' free apps', 'help teachers', and 'enormous data' were used by the respondents to indicate the usefulness of coronas corpus in the classroom.

2.1.1b Theme: Increased knowledge of the subject matter

Based on the findings, 3 respondents connected the use of corpora in language and knowledge acquisition. R4 said that "corpus-based activities have increased my guessing of the meaning from the content of the texts" (R4). Followed by " Increases knowledge and understanding variation of syntax" (R5). Finally, further agreed with another respondent who said that "it helps in detecting important words"(R7).

Another constraint that the respondent faced at the beginning of using corpora was the difficulty of determining to get the subject matter due to a large number of texts. However, as time goes they were all able to do you with much ease. One respondent said, ". .at the beginning difficult to determine the subject. But, I managed to understand very well." (R5). Another respondent agreed, '. .No. It was easy once the hands-on activity started"(R7). Only 1 respondent said that it was entirely difficult from the beginning. This respondent said, ". .yes. Somehow we have to determine what we want to teach" (R4).

Based on ye responses words like 'increased meaning guessing", "increases knowledge understand variation of syntax" and " helps detect words" which derive from the theme above.

2.1.1c Theme: Tool for Innovation

Based on the findings, most respondents felt that using corpora in teaching and learning has a positive effect. One respondent said, " it will be the innovation, so will attract the students' (R2).

2.1.1d Theme: Tool for big language analyse data

Further points given by another respondent are that " The positive aspects is easier to analyse the data" (R3). Another respondent said, " useful tool for compiling and analysing language data on the spot. " (R2).

2.1.1e Theme: Easily integrated into the syllabus

In addition to this, another respondent said that "corpus can be integrated in school syllabus to enhance language learning" (R4).

2.1.1f Theme: Ease in creating authentic and variety of activities

Finally, another respondent said, "you can use corpus to create various activities" (R6). Which was supported by other respondents who said, ".enlighting becos the materials are real life and according to the cultural context."(R2).

Overall, the findings of the benefits aspects were supported by the following studies. There are numerous advantages to using the CiDDL-Pd technique. It emphasises the significance of lexis and grammar (lexicogrammar) in ESP, as well as specialised phraseology. Grammar and vocabulary are viewed as 'complementary' units by neo-Firthians (Yunus, 2017). Furthermore, the CiDDL-Pd method improves language learning by utilising a variety of strategies. Rather than using the unauthentic teaching aids in academic books, settings might be created using technology. Additionally, learners have the opportunity to extend their pattern memory retention (Cobb, 1997; Nation, 2000). Where research into the use of corpora in language pedagogy is concerned, various research (Al-Thunaibat & Singh, 2020; Hajimia et al., 2019, 2022; Noorizan, Shamsudin, Manan & Singh, 2017) have been conducted to show the ultimate benefits of such an approach.

2.2 Disadvantages
2.2.2a Theme: Technical constraints

Based on the findings, all the respondents said that the weakness of the internet caused problems in creating activities using corpora. One respondent said that 'the system not stable' (R2). Another respondent said that 'WiFi too slow (R3). Finally, respondents said that, internet slow and the texts in the corpus are difficult to analyse due to this' (R5).

Based on the response, words like 'system not stable', 'slow wifi', ' slow internet' and 'difficult to analyse' were retrieved resulting in the theme above.

2.2.2b Theme: Not adaptable to foreign writing systems and forms

A few respondents said the corpora is difficult to be used in foreign languages with characters such as the Chinese and Arabic languages known and hieroglyphs and logograms. For example, one respondent said, "…cannot analyse the data for other languages such as Chinese character…" (R3). This was further agreed by another respondent who said, " …I can set commands of Arabic characters in the software, But when actually analyzing its is difficult for the system to understand the actual arrangement of the word…it was more of a direct transcription then actual word order…(R7). This was supported by another respondent who said, "…it might be suitable for certain language skills and not for all skills in language acquisition." (R6).

2.2.2c Theme: Relevance of prior knowledge

Two respondents raised the issue of background knowledge of the language's grammatical orders and patterns is essential in determining the accurate usage of the language. One respondent said, "…corpus will never tell you if a sentence is grammatically or syntactically plausible or not…"(R1). This statement was further agreed upon by another respondent who said, "…we have no way of distinguishing whether there are mistakes or not…unless one of us is the expert…"(R7).

Overall, the findings of the negative aspects were supported by the following studies. However, CiDDL-Pd has also been found to have some limitations. Yunus (2017) listed issues with practicality, as accepting a new teaching and learning approach may not be as soundly received by many. Other issues include "different learners' learning styles, technophobic students, teachers' resentment of new technology, and some other barriers for instance technical and logistic aspects" and the difficulty for the tool to adapt to foreign languages (Johns, as cited in Bada, 2015, p.69). These limitations have impacted the research interest and adaptation of CiDDL-Pd in educational settings even though many CiDDL-Pd resources can be found online. Nevertheless, CiDDL-Pd has been profound as a novel and effective approach. CiDDL-Pd is a "ground-breaking approach that allows teaching and learning in various ways" (Rapti, 2013, p.17), particularly in ESP context. CiDDL-Pd facilitates ESP educators to create materials for language teaching pedagogy as well as for ESP courses (Chambers & Kelly, 2004).

Q3 Is there a significant difference between Malaysian educator's perspectives compared to other scholars

Overall Malaysian lecturer's perspective about the application of CiDDL in the classroom seems to have positive significance with scholars from other countries. This data was collected based on the observation. Three aspects were agreed upon, which was similar to most others scholars, namely; teaching methodology, language content and using corpus tool.

Teaching Methodoloy

The educator's had agreed during the discussions that learners should be taught the full norms of usage for the words "but" and "however". The issue of how to implement this in the classroom has to be addressed in their writings. In general, the educators agreed that it was crucial to give beginners in particular clear and trustworthy norms.

Here, it's important to note the presumptions that the participants have regarding their prospective students. The educators frequently drew on their prior academic learning experiences throughout the class discussion. Some educators were sharing their prior knowledge of students' backgrounds and their own teaching experiences because they believed it was crucial to create a safe learning environment to students. To do this, they taught clearly defined rules, provided information that was simple to understand, and avoided overwhelming students with linguistic complexities.

The topic of authentic texts and vocabulary difficulties came up frequently. Different approaches to dealing with actual texts in the classroom were investigated, such as changing the task rather than the text.

Language content
The educators initially focused on the immediate issue at hand—how to teach the use of "but" and "however"—but this rapidly led to a larger discussion on crucial questions like

what language to teach. As a result, they became more critical of the material in the textbooks.

The activity made the educators aware of the challenges that can come with teaching real language use, even if they largely agreed that it was important. This is apparent in the tension they exhibit between wanting to educate novices effectively but without sacrificing teaching the actual grammar rules of English.

Corpus tool

The educators understood the significance of the corpus and corpus tools as resources for the learner to investigate the complexity of language as well as to offer credibility by enabling the learner to investigate real texts and discover language use at their own speed. They also seem to think that after learning the condensed rules for the linguistic aspect, the corpus tools will assist the students figure out the full set of rules on their own.

Conclusion

The Malaysian Corpus of Online Newspaper Articles (MyCORONAS) corpus, which was developed by the researcher for her doctoral thesis, was used, as well as a few corpus tools. The qualitative data was analysed using thematic analysis as proposed by Braun & Clarke (2016). This research yielded two results: the first being the successful application of a CiDDL-Pd framework for teaching language in classrooms (CiDDL-TG), and the second being the development of a collection of CiDDL-Pd based lesson plans. The significant output, as intended, would result in a good enhancement in teachers' perception of the corpus' application in classroom activities while teaching. The outcome of these findings is a framework indicating the benefits and disadvantages of using My Coronas Corpus in teaching and learning shown in Figure 1 below.
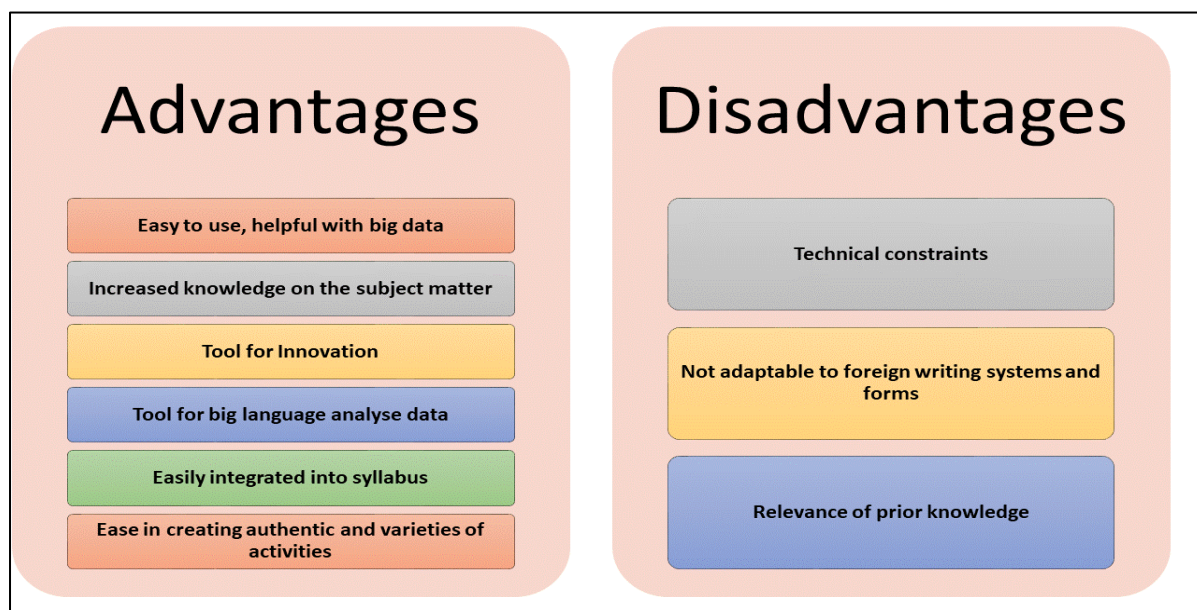

Figure 1. *Advantages and disadvantages of applying CiDDL-Pd using MyCoronas Corpus*

In conclusion, the present study found that there is a more positive impact of CiDDL-Pd on language teaching and learning. Using a corpus-based approach seems to contribute significantly to understanding how language works. With ample resources available for the development of corpora, the next thing teachers must consider and work on will be to bring the

corpora into the classroom. One such way will be by applying the proposed CiDDL-Pd teaching method in classrooms.

References

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. Journal of Research & Method in Education, 5(6), 66-70.

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. Journal of Research & Method in Education, 5(6), 66-70.

Benavides, C. (2015). Using a corpus in a 300-level Spanish grammar course. Foreign Language Annals, 48(2), 218–235.

Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. ReCALL, 19(3), 307–328.

Breyer, Y. A. (2011). Corpora in Language Teaching and Learning: Potential, Evaluation, Challenges. English Corpus Linguistics. Volume 13. Peter Lang GmbH, Internationaler Verlag der Wissenschaften. Eschborner Landstrasse 42-50, D-60489 Frankurt am Main, Germany.

Breyer, Y. A. (2011). Corpora in Language Teaching and Learning: Potential, Evaluation, Challenges. English Corpus Linguistics. Volume 13. Peter Lang GmbH, Internationaler Verlag der Wissenschaften. Eschborner Landstrasse 42-50, D-60489 Frankurt am Main, Germany.

Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. Language Teaching, 52(4), 460–475.

Corino, Elisa, Christina Onesti (2019). "Data-Driven Learning: A scaffolding methodology for CLIL and LSP teaching and learning". Frontiers in Education 4/7. DOI: 10.3389/feduc.2019.00007

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. Computer Assisted Language Learning, 28(2), 130–144.

Forti, L and Spina, S (2019) "Corpora for Linguists vs. Corpora for Learners Bridging the Gap in Italian L2 Learning and Teaching". EL.LE Vol. 8 – Num. 2 – https://doi.org/10.30687/elle/2280-6792/2019/02/005.

Hajimia, H. (2018). A Corpus-Based Analysis on the Functions of Discourse Markers used in Malaysian Online Newspaper Articles. International Journal of Research in Humanities and Social Studies, 5(1).

Hajimia, H., Nordin, N. R. M., Singh, M. K. S., & Golingai, P. (2022) The rhetorical structure of newspaper reports: A synergy between corpus, genre and discourse analysis. Indonesian Journal of Applied Linguistics, 12(1), 88-101.

Hajimia, H., Singh, M. K. S., & Chethiyar, S. D. M. (2020). Second Language Acquisition: Krashen'S Monitor Model and the Natural Approach. International Journal of Social Sciences, 6(3), 87-99.

Hajimia, H., Singh, M. K. S., & Rathakrishnan, M.(2019). Analyzing Frequency And Distributional Pattern Of Discourse Markers In The Malaysian Newspaper. Opción, Año 34, Especial No.19 :1692-1714 ISSN 1012-1587/ISSNe: 2477-9385

Hajimia, H., Singh, M. K. S., Yusoff, M. D. D. N., & Rathakrishnan, M. (2019). English Grammar Revolution using Corpus-Based Approach: How do the Students React?. Strength for Today and Bright Hope for Tomorrow Volume 19: 10 October 2019 ISSN 1930-2940, 232.

Jamal, J., Shafqat,A., Afzal,E (2021) Teachers' perceptions of incorporation of corpus-based approach in English language teaching classrooms in Karachi, Pakistan, Liberal Arts & Social Sciences International Journal (LASSIJ) Vol. 5, No. 1, (January-June 2021): 611-629 Research Article | Original Research https://doi.org/10.47264/idea.lassij/5.1.40

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The sketch engine: Ten years on. Lexicography, 1, 7–36.

Leńko-Szymańska, A. (2014). A teacher-training course on the use of corpora in language education: Perspectives of the students. Insights into Technology-enhanced Language Pedagogy. Peter Lang: Frankfurt am Main, 129-144.

Leńko-Szymańska, A. (2014). A teacher-training course on the use of corpora in language education: Perspectives of the students. Insights into Technology-enhanced Language Pedagogy. Peter Lang: Frankfurt am Main, 129-144.

Qiong, H. (2007). Corpora and English teaching: Retrospect and prospect. International Journal of Scientific Research and Innovative Technology, 4(7), 65-73.

Qiong, H. (2007). Corpora and English teaching: Retrospect and prospect. International Journal of Scientific Research and Innovative Technology, 4(7), 65-73.

Rapti, N. (2013). Data-driven grammar teaching and adolescent EFL learners in Greece. Studies in Variation, Contacts and Change in English. Corpus Linguistics and Variation in English: Focus on Non-Native Englishes.

Römer, U. (2011). Corpus research applications in second language teaching. Annual Review of Applied Linguistics, 31, 205–225.

Singh, M. K. S., & Shamsudin, S. (2010, May). Corpus linguistics: Syntactical analysis of learner corpus anyone. In the proceedings of The 7th International 'Language for Specific Purposes'(LSP) Seminar on 'Glocalisation of New Literacies'. Universiti Teknologi Malaysia, Johor. Malaysia.

Singh, M. K. S., & Shamsudin, S. (2010, May). Corpus linguistics: Syntactical analysis of learner corpus anyone. In the proceedings of The 7th International 'Language for Specific Purposes'(LSP) Seminar on 'Glocalisation of New Literacies'. Universiti Teknologi Malaysia, Johor. Malaysia.

Szudarski, P. (2018). Corpus linguistics for vocabulary: A guide for research. New York, NY: Routledge.

Vyatkina, Nina (2020). Corpora as open educational resources for language teaching. Foreign Language Annals. doi:10.1111/flan.12464.

MAILING INFORMATION
Hafizah Hajimia
Department of English and Linguistics, Academy of Language Studies,
Universiti Teknologi MARA, Perlis, Malaysia
hafizah.hajimia@uitm.edu.my
ORCID ID: https://orcid.org/0000-0002-9235-9357
Manvender Kaur Sarjit Singh
School of Languages, Civilization and Philosophy
Universiti Utara Malaysia
manvender@uum.edu.my
Nor Azira Mohd Radzi
Academy of Language Studies
Universiti Teknologi Mara Cawangan Perlis
norazira202@uitm.edu.my
ORCID ID: 0000-0002-6757-5226
Nur Rasyidah Mohd Nordin

School of Languages, Civilization and Philosophy
Universiti Utara Malaysia
nurrasyidah@uum.edu.my
ORCID ID: 0000-0001-7954-0250
Nur Asyrani Che Ismail
Academy of Language Studies
Universiti Teknologi Mara Cawangan Shah Alam
nurasyrani@uitm.edu.my