TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

BY

NAJAM IRSHAD

NATIONAL UNIVERSITY OF MODERN LANGUAGES

ISLAMABAD

September, 2024

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

By

NAJAM IRSHAD

B.S. English, Abbottabad University of Science & Technology, 2020

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF PHILOSOPHY

In **English**

To

FACULTY OF ARTS & HUMANITIES

NATIONAL UNIVERSITY OF MODERN LANGUAGES, ISLAMABAD

FACULTY OF ARTS & HUMANITIES
NATIONAL UNIVERSITY OF MODERN LANGUAGES

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with the overall exam performance, and recommend the thesis to the Faculty of Arts & Humanities for acceptance:

Thesis Title: Teaching Discourse Markers at Higher Secondary School Level Using Data-Driven Learning Approach: An Experimental Study

Submitted By: Najam Irshad **Registration** 90 MPhil/Eng Ling/S21

Dr. Azhar Habib _____
Name of Supervisor Signature of Supervisor

Dr. Farheen Ahmed Hashmi _____
Name of Head (GS) Signature of Head (GS)

Prof. Dr. Arshad Mahmood _____
Name of Dean (FAH) Signature of Dean (FAH)

Prof. Dr Muhammad Safeer Awan _____
(Name of Pro- Rector Academics) Signature of Pro-Rector Academics

_____
Date

AUTHOR'S DECLARATION

I, Najam Irshad

Son of Muhammad Irshad

Registration # 90 MPhil/Eng Ling/S21

Discipline English Linguistics

Candidate of **Master of Philosophy** at the National University of Modern Languages do hereby declare that the thesis **Teaching Discourse Markers at Higher Secondary School Level Using Data-Driven Learning Approach: An Experimental Study** submitted by me in partial fulfillment of MPhil degree, is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in future, be submitted by me for obtaining any other degree from this or any other university or institution.

I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be cancelled and the degree revoked.

_____

Signature of Candidate

_____

Name of Candidate

_____

Date

ABSTRACT

Title: Teaching Discourse Markers at Higher Secondary School Level Using Data Driven Learning Approach: An Experimental Study

The complexities in defining language concepts persist despite extensive research, something which is also reflected in language teaching. It has, consequently, continued to fall behind. The subject study, experimental in nature, was conducted to see if Data Driven Learning could be an effective teaching methodology for linguistic features which are inconsistent in origins and uses, which in this study are Discourse Markers (DMs). The study sample comprised 90 learners in 12th grade from a school in Abbottabad, divided into 3 groups of 30 each: a control group, treated with the Presentation-Practice-Production (PPP) method; and two experimental groups, each of them treated with hard DDL (computer-based) and soft DDL (paper-based). They were tested using an essay writing exercise and a sentence making exercise, with the objective to find out if the two versions of DDL had higher learning outcomes than the PPP method as well as how they compare with each other. A 3-month delayed post-test was also employed to gauge long-term retention. The theoretical framework forming the basis for the study comprised the Data Driven Learning Approach, the Noticing Hypothesis, Input Enhancement and the Involvement Load Hypothesis. The post-test results were comparable for both versions of DDL in using DMs in diverse ways and applying them in actual text; significantly higher than the control group. All three groups showed a decline in the delayed post-test, however, it was sharper for the SEG and the control group compared to the HEG. Hard DDL demonstrated higher long-term retention, while soft DDL was only slightly better at best. The results imply that Hard DDL has higher involvement load, noticing and cognitive stimulation, leading to a more solid foundation of concepts, while Soft DDL does not have significant benefits over a long time. The results provide grounds Hard DDL can be a suitable and more effective replacement for teaching Discourse Markers at higher secondary level in Pakistan. This research also opens the door to explore the differences in performance of the two different versions of DDL.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

DDL: Data Driven Learning

DM: Discourse Marker

PPP: Presentation-Practice-Production

SEG: Soft Experimental Group

HEG: Hard Experimental Group

ACKNOWLEDGEMENTS

I am grateful to Allah Almighty for blessing me so that I could complete this thesis which was an arduous task.

I am eternally thankful to my supervisor Dr. Azhar Habib for his constant support and guidance. I am also grateful to the Dean, Faculty of Arts and Humanities, Prof. Dr. Arshad Mahmood, Head of English Graduate Studies, Dr. Farheen Ahmed Hashmi, and all the respected faculty for helping me through-out this process.

I would like to take this moment to thank and appreciate my family who were a constant source of support and morale in my journey. My friends also played a huge part in pushing me through this process and helped me with their encouragement, kind words as well as other methods of support, for which I will forever stay obliged.

Thank you all!

DEDICATION

I dedicate this thesis to my amazing parents, sweet siblings and my incredible friends, who remained my light through this journey and without whom it would not have been possible.

CHAPTER 1

INTRODUCTION

1.1 Overview

This section illustrates key concepts, introduces the objectives and the hypotheses and presents the significance of this current study.

1.2 Discourse Markers

Discourse markers are a class of linguistic items that play a role in organizing discourse (Maschler & Schiffrin, 2015). There is disagreement on the exact definition of discourse markers and they have been defined differently by many different researchers. Redeker (1991, cited in Sun, 2013) defines discourse markers as "Linguistic expressions that is [are] used to signal the relation of an utterance to the immediate context with the primary function of bringing to listener's attention a particular kind of the upcoming utterance with the immediate discourse context" (p. 2136). Schiffrin (1987) offers another definition, defining discourse markers as "dependent elements which bracket units of talk" (p. 31). Jones and Carter (2014) believe that the purpose of DMs is to make the discourse coherent. From the above definitions, it can be agreed that the basic function of discourse markers is the organization of text. It performs the function of connecting the preceding sentence with the following sentence. Examples of discourse markers include the bold words in the following sentences:

a. He is a good person. **However**, fate did not favor him.

b. Saad was not picking up his phone. **So**, I could not invite him to the party.

c. I hate him too. **But** still, we have to take care of him, because he is part of the family.

d. The team left very soon. **Furthermore**, they left a confidential message for the message.

e. **You see**, she is different than the lot.

f. **Listen**, you'll never be forgiven.

g. **I think** I might not be able to make it.

h. **Oh,** it's a shame it failed before even starting.

From the above examples, we can make a few observations about discourse markers. The first is that discourse markers can belong to different parts of speech. Secondly, discourse markers may not always carry meaning for the sentence, like in the case of "you see" and "oh". Their function is mostly pragmatic. Maschler and Schiffrin (2015), in fact, divide them into three categories based on their function: textual, interpersonal, and cognitive.

There are a few difficulties in teaching discourse markers. One is that they are not from one lexical group. Therefore, teachers have to account for different forms of words. The second problem is that the meaning of discourse markers varies according to context. They perform the same function at different intensities depending on the situation. Therefore, teaching a specific function of a discourse marker through traditional means could be a hindrance to the understanding of all the different meanings of one.

1.2.1 Why Teach Discourse Markers

Discourse markers are considered very important for communication and to properly convey a message. They allow language users to produce coherent and well-organized text (Ying, 2007). Haberlandt (1982) believes that discourse markers reduce the reading time for a text. Fraser (1993) propositioned that discourse markers indicate the speaker's perception and awareness about the connection between subsequent sentences. They help make the proper flow of the discourse by setting and connecting the sentences. Their exclusion does not make a sentence ungrammatical, however, without them; the discourse would look disorganized and unnatural.

Much research has gone into the role discourse markers play in a discourse. Stenstrom (1994) considers them textual devices that organize and guide the discourse as well as are used to create boundaries within discourse. Redeker (1991) maintains that the primary function of discourse markers is to link an utterance with its context in the discourse. They are used in terms of their procedural meanings, instead of conceptual meanings, which can change according to each situation. Thus, they indicate the role of a speaker's utterance in the discourse.

The importance of discourse markers is often found in coherence-based models, which consider discourse markers essential for coherence in text. Schiffrin (1987) considers relationship markers between the sentences of discourse in terms of their syntactic and semantic properties. They serve as "contextual coordinates for utterances by locating them on one or more planes of talk" (Fung, 2003, p. 46) essentially serving as a discourse glue (Aidinlou & Mehr, 2012; Alraddadi, 2016; Rahimi & Riasati, 2012). They indicate the context of the utterance both, in terms of the speaker and the preceding and following discourse. Redeker (1991), in her model of coherence, provides three components responsible for coherence: ideational structure, rhetorical structure, and sequential structure. Ideational structures indicate the speaker's commitment to the utterance. Rhetorical structures represent the relationship between the intentions of the utterances. Sequential structures allow the transition of one thought into another as well as topics. She believes that discourse directly contributes to the above structures. Maschler and Schiffrin's (2015) differentiation of discourse markers into referential, structural, and cognitive functions aligns with the coherence structures described by her, thus making discourse markers an essential part of language use and consequently warranting explicit teaching in language classrooms, which has been proved by many studies to be effective in improving use (e.g. Alraddadi, 2016; Aidinlou & Mehr, 2012; Rahimi & Riasati, 2012)

1.3 Corpus Linguistics and English Language Teaching

Corpus is a Latin word that means "body". In the context of modern Corpus Linguistics, it is quite literally a body of text. According to Crystal (2011), a corpus is "a collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting-point of linguistic description or as a means of verifying hypotheses about a language" (p. 117). It involves large collections of language, which are computerized databases created for linguistic research. To answer the question of what corpus linguistics entails, Kennedy (2014) describes corpus linguistics as a field that encompasses the process of collecting spoken and written texts and analyzing them to extract evidence to describe the nature, structure, and use of languages.

1.3.1 Properties of Corpora

Corpora are generally associated with the following properties (Bennett, 2010; McEnery & Wilson, 2001):

1. Corpora can never produce a negative result.
2. They use samples of natural language; it may involve both, speech and text corpuses.
3. They are representative of the whole language or the genre they belong to.
4. They are stored electronically.

1.3.2 Use of Corpora in Language Teaching

The tendency to rely on corpus-based evidence has been increasing more and more in the last couple of decades (Lüdeling & Kytö, 2009). Corpus-informed research is becoming more and more prevalent, which is used to make a lot of decisions in the real world. Language teaching is one of the biggest real world applications of corpus. Corpus techniques have enormous potential to benefit educational decision making as well as the actual classroom. The biggest strength of using a corpus is its empirical nature, which provides solid evidence for teachers and curriculum developers to make decisions about what they should teach in their classes. A corpus usually consists of a large number of texts that reflect natural language use and its analysis can offer many insights based on the patterns of lexis, grammar, semantics, pragmatics, and textual features.

Corpora can be used in two ways for language teaching:

1. Indirect use of corpora
2. Direct use of corpora

1.3.2.1 Indirect use of corpora

Indirect usage involves the use of corpora in which the learner does not directly engage with the corpora. Rather, corpora are used to form and inform the several resources and materials that aid in learning a second language. In addition to that, the direct use of corpora faces several challenges, that limit its use, such as the inexperience of the learners and teachers in using corpus tools, computer illiteracy

and the unavailability of technology to access corpus tools. However, the indirect use of corpus tools, i.e. by those who develop and publish educational material, serves to inform classroom learning. The various forms of indirect use are discussed below:

1.3.2.1.1 Reference Publishing/lexicography

Dictionaries serve as excellent reference material. The advancements in corpus linguistics have revolutionized lexicography, which now provides tools such as parts of speech tagging as well as semantic and syntactic tagging. These tools can analyze vast bodies of texts in a very short time and generate word lists, and provide their contexts of appearance as well as how they behave in different contexts. They allow lexicographers to extract all the examples of typical, authentic usage of words in little time. It allows them to examine the different ways that words behave, what their features are and how they collocate (Lew & Jackson, 2013).

For language teaching, the most important use of corpus tools is the development of learner dictionaries. The use of authentic examples has greatly innovated dictionaries and referencing material because learners can get a better idea of how different vocabulary is used. Corpus tools can also provide information about how words collocate and how idioms and pragmatic expressions behave. These tools are especially useful for learners of English for a specific purpose (ESP). Learners can often get overwhelmed by a large number of vocabulary that is irrelevant to their purposes, therefore, a learner dictionary made for specific purposes, which not only has vocabulary related to their required field, but also has definitions that are most accurate for the target environment (Nurmukhamedov, 2012). Moreover, corpus-based reference material, such as word lists and frequency data, can be frequently updated.

1.3.2.1.2 Syllabus Design and Material Development

Not only has the use of corpora allowed more accurate and authentic descriptions of language use, it also provides evidence to inform the decisions about what to teach in a classroom and what should be included in teaching material. Mindt (1996) points out most English textbooks seem to teach a version of the English language that only exists in the classroom. It does not represent how English is actually used in real-world scenarios. Corpora provide more realistic examples of language usage that reflect the complexities and nuances of natural language. In

addition to this, corpora allow the incorporation of collocations in teaching material. Teaching collocations is an important part of language learning because they facilitate communication. Language consists of a lot of fixed phrases that are frequently used. Thus, finding out and teaching these phrases can significantly improve a learner's communication in a second language (Rao, 2018).

1.3.2.2 Direct Use of Corpora

The direct use of corpora is related to its use in language classrooms. According to Meyer (2002), direct use in classrooms takes a more exploratory approach, in which the teacher and learner find out patterns and information about the language. Direct use of corpora usually occurs in more advanced classrooms, due to the reasons discussed earlier, such as the requirement of computer literacy and the ability to use corpus software.

1.3.2.2.1 Data Driven Learning

Johns (1991a, 1991b, 2002) coined the term 'Data Driven Learning' (henceforth 'DDL'), which describes a learner-centered approach to language teaching, which shifts the teacher's role to a facilitator in the class. It involves the use of a 'concordancer', a computer program that is used to display all the instances of occurrence of a piece of language in a corpus.

In this method, learners try to discover the answers to language questions using corpus tools. In this process, incomplete and partial generalizations are extracted from the corpora as a way to eventually build up to a fully satisfactory rule about language. Since this is a bottom-up approach, it relies on the induction of the rules to the learner, where they observe the patterns discerned by corpus tools and form their own hypothesis about language rules.

One of the main advantages brought by corpora is the difference between consulting and learning. While one might think that the act of looking up a word in a dictionary (consulting) might lead to longer retention, this is not the case. Consulting is usually less long-lasting than when something is learned by a learner him/herself (Boulton & Landure, 2016). The act of learning involves looking up a word in its context and finding out the patterns of its use. This can occur in two ways: teacher-directed and learner-lead. In the teacher-directed method, the teacher highlights the

errors made by a learner and provides examples from a corpus accordingly. For example, if a learner uses the wrong preposition with a word, the teacher can provide links to a relevant concordance, which provides examples of the prepositions that usually occur with a word, thus allowing the learner to learn the correct preposition. Similarly, the teacher can highlight words that usually do not collocate together and provide examples from the corpus of the usual collocations of that word. In the learner-led method, the teacher is not involved at all. The role of the teacher is reactive. Learners independently explore their findings using different corpus tools and the teacher's role is to provide feedback on them. This leads to even more inductive learning, which is longer lasting and tends to be better understood (Boulton & Landure, 2015).

1.3.2.2.2 Types of Data-Driven Learning

Leech (1997) describes two versions of DDL that could be implemented in class: **the soft version and the hard version**. According to Gabrielatos (2005), the hard version strictly adheres to corpus software and is more hands-on. Learners directly engage with the corpus using CALL (computer-assisted language learning) software and freely explore the corpus. The implementation of the hard version is difficult, because it requires more resources, as the whole class needs access to computers. The soft version, on the other hand, can be implemented with fewer resources. Computer access is only required for the teacher. The teacher can manipulate the corpus and draw out examples that are suitable for the class. These examples then can be provided to the learners in the form of handouts for them to exploit and use for data-driven learning.

With increased interest and wider access to corpus technology, Data-Driven Learning (DDL) has also expanded its horizons and is no longer limited to inductive learning. Deductive approaches to DDL have also been developed, in which the learners first infer the rules, after which they are presented with further explanations and examples from the instructor. Teachers can also provide expanded context to facilitate better learning and understanding of the patterns of the language (Chujo, Anthony & Oghigian, 2009, cited in Shaw, 2011). The varying approaches to DDL still keep its essence, which is to expose learners to as much authentic language data as possible. Fraser (1996a, 1999, 2010) also discusses the benefits of providing

different methods of exploring language data, including providing more contextualization. She also points to another important feature of later uses of corpus data, which is the focus on linguistic forms that are not frequently used by language learners. In this way, learners can learn what kind of unnatural linguistic forms native speakers produce and have a better understanding of the patterns and rules that language follows.

1.4 The Statement of Problem

The ability to use a language is far more than simply forming grammatically correct sentences. Effective communication demands that not only do these sentences need to have connections with each other, but they must also be relevant to the immediate context surrounding the language use. Discourse markers (DMs) are among the primary linguistic elements responsible for this purpose. However, DMs present a certain level of difficulty in their teaching and use, because their uses and origins vary a lot. This is evident by the large number of studies available finding lower levels of DM use in non-native speakers. There have been many attempts to bridge this gap; however, most have not seen much success. Data Driven Learning (DDL), a corpus-based approach, here presents a new possibility of filling in the void, its inductive and interactive nature may be more effective in inducing learners to acquire these features of language in a manner similar to the acquisition of first language. This is an alternative to learning DMs as linguistic items and rules, which rely on linguistic description, which, no matter how extensive, can never be without fault or deficiency. This older method is reflected in the form of the Presentation-Practice-Production (PPP) method traditionally employed in classrooms. There has been some research in testing DDL in teaching linguistic items with various levels of success. However, there are two areas where the research lacks. The first such area is research in lower-level learners. Most of the research targets university students or learners in advance language courses. Secondly, there has been very little work in exploring how Pakistani learners respond to the DDL approach, with very few studies available on public sources. As such, this study is positioned perfectly to fall into place to fill these gaps. Thus, this study sought to explore how the two versions of DDL, a paper-based soft version and a computer-based hard version, compare to the PPP method in enabling learners to acquire the ability to use DMs. The study also

goes further into measuring their long-term retention to ensure that there are material benefits of choosing one method over the other in the long run.

1.5 Objectives of the Study

This section outlines the objectives of this research study. These objectives were used to guide all the decisions about this study ranging from the theoretical framework to methodology to the analysis of the data collected during it.

- To determine whether the different versions of Data-Driven Learning yield better results than the Presentation-Practice-Production method
- To determine whether the soft Data Driven Learning produces similar results in comparison with the hard Data Driven Learning.
- To find out if different versions of Data Driven Learning have better long term retention of ability in comparison with the Presentation-Practice-Production method.

1.6 Hypotheses

1.6.1 Null Hypotheses

1. Null Hypothesis 1: There is no significant difference in the post-test scores of the three groups
2. Null Hypothesis 2: There is no significant difference in the delayed post-test scores of the three groups
3. Null Hypothesis 3: There is no significant difference in the pre and post-test scores of the Control Group
4. Null Hypothesis 4: There is no significant difference in the pre and post-test scores of the Soft Experimental Group
5. Null Hypothesis 5: There is no significant difference in the pre and post-test scores of the Hard Experimental Group
6. Null Hypothesis 6: There is no significant difference in the post-test and delayed post-test scores of the Control Group
7. Null Hypothesis 7: There is no significant difference in the post-test and delayed post-test scores of the Soft Experimental Group

8. Null Hypothesis 8: There is no significant difference in the post-test and delayed post-test scores of the Hard Experimental Group

1.6.2 Main Hypotheses

1. Hypothesis 1: Learners learning through the hard version of the Data-Driven Learning Method show a higher level of improvement in their use of discourse markers in comparison with the PPP method
2. Hypothesis 2: Learners learning through the soft version of the Data-Driven Learning Method show higher level improvement in their use of discourse markers in comparison with the PPP method
3. Hypothesis 3: Both the soft and the hard versions of Data-Driven Learning have similar results in improving learner proficiency in using Discourse markers in the post-test.
4. Hypothesis 4: Teaching through both versions of Data-Driven Learning results in a higher level of long-term retention in comparison with the PPP method.

1.7 Significance of the Study

The current study finds its uniqueness in several ways. The most basic point of differentiation is its placement in Pakistan where research on Data Driven learning as well as the teaching of Discourse Markers has been scarce with very few studies available on these topics. Aside from that, the current study targets higher secondary school students, a demographic that has been mostly neglected in DDL studies due to the assumption that it may only be suitable for higher level learners as well as due to the difficulty of arranging the appropriate equipment in the past. This is evident in Bao's (2021) meta study on research on DDL where the participants of most of the quoted studies were university students. Research has also focused mostly on oral discourse markers and written discourse markers have found lower attention (e.g. Aijmer, 1996; Redeker, 1991; Schiffrin, 1987), leading to a gap in the research.

In addition to the above, this research is significant in three different ways. First, it ventures to find a solution to the lower levels of proficiency in using discourse markers. Direct exploration of language data presents itself as a possible way to overcome the difficulty in teaching discourse markers owed to the limitations of

teaching meaning and usage through a PPP method. Secondly, this research will add to empirical knowledge about the difference in achievement through soft and hard versions of Data-Driven Learning, which has mostly seen theoretical and qualitative work. Finally, this research will deal with the implementation of Data-Driven Learning, both hard and soft versions, on higher secondary levels; a level where DDL has not been much explored, because it was previously assumed that only higher-level learners can use corpus software.

1.8 Delimitations

This research study was delimited to a high income school in Abbottabad. The subjects of the study were learners enrolled in 12th grade. The study was conducted in a period of 6 months.

1.9 Summary

This chapter introduced the key concepts in the study, established the objectives of the study, introduced the hypotheses and discussed how the study finds its place in the currently available knowledge. It established DMs as essential parts of language use and provided evidence in their scarcity in the discourse of Pakistani English language users. It further introduced Data Driven Learning as an alternate method of language teaching which seems promising as the solution to solving the problem of lower proficiency. It carved out a space for the current study to take place with its unique attempt to explore a DDL based approach to teaching DMs in Pakistani higher secondary schools.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

The purpose of this chapter is to delve into the existing literature about the various variables involved in this study. It defines key concepts in light of available works, looks at past developments, reviews methodology and comments on how they add to the current body of knowledge.

2.2 Defining Discourse Markers

The question that any discussion of discourse markers first asks is what discourse is, and this is exactly the question that Deborah Schiffrin (1987) opens with in her book. In discussing discourse through a lens of discourse analysis, she puts forth some key assumptions about language or discourse. These assumptions provide some key characteristics of discourse, the first being that language always occurs in context, its comprehension is shaped by such context, and it is used and designed for communicative purposes. These assumptions for language provide us with the basis for the existence of discourse markers. As Schiffrin (1987) herself points out, language exists in both grammatical and cognitive contexts and any attempt to produce language without these will prove to be useless. The point to note here is that if language does indeed exist in context, there must be something within it connecting it with it on both syntactic and semantic levels. This gives us a starting point for our discussion about discourse markers, which are considered to perform the various functions that are required in making language contextual and communicative.

Many terms have been used to define structures in language that are generally referred to as discourse markers. Various linguists have tried to define their own terms for them including discourse operators (Redeker, 1991), discourse particles (Schourup, 1999), pragmatic particles (Östman, 1981), pragmatic markers (Fraser, 1999), or connectives (Blakemore, 1987) among many others. The wide range of terms used to describe DMs is an indicator of the broad number of concepts and functions attached to discourse markers. Schiffrin (1987), in her operational definition of DMs, calls them "**sequentially dependent [emphasis original]** elements which bracket *units of talk* [added emphasis]" (p. 31). The use of the term 'units of talk' is a

careful choice by her, because there are so many varied elements that influence the use of DMs that using any other specific term will limit the scope of any discussion of DMs, which are vast in actuality.

This study will use the term Discourse Markers, as it covers the most number of meanings and is the least contentious one (e.g. Jucker, 1998; Schourup, 1999). They perform various functions such as indicating turns, connecting discourse to the context, signaling relationships, indicating intimacy, and conveying attitude and relative standing to topics of conversation (Jucker, 1998). DMs have been defined in different ways by different researchers. The definitions each come with their own concept of the function that DMs perform in text, as well as their linguistic nature. The struggle to reach a consensus on the definition of DMs continues, despite wide research into their usefulness, function and forms in the English language. This partially stems from the fact that DMs belong to various linguistic categories including conjunctions, interjections, adverbs, etc (Redeker, 1991). The second reason is the disagreement among linguists about the different aspects of discourse markers. Redeker's (1991) definition of DMs describes them as "a word or phrase — for instance, a conjunction, adverbial, comment clause, interjection — that is uttered with the primary function of bringing to the listener's attention a particular kind of linkage of the upcoming utterance with the immediate discourse context" (p.30). The most important thing to note here is Redeker's use of the word "utterance" here, effectively limiting DMs to oral use. Moreover, it restricts the function of discourse markers to only organizing utterances with the rest of the discourse and ignores other functions of discourse markers as we have previously mentioned. Redeker (1991) in the same documents cites Schiffrin's (1987) definition of DMs, stating that "Discourse markers are linguistic, paralinguistic, or nonverbal elements that signal relations between units of talk by the virtue of their semantic and syntactic properties (if any) and, most importantly, by the virtue of their sequential position as initial or terminal brackets demarcating discourse unit". A particular strength of Schiffrin's definition is her recognition of the relatively less defined nature of DMs in language as well as their existence as non-linguistic elements. However, this definition, too, has its weaknesses. It only considers the function of DMs on a structural level and ignores their interpersonal functions (Jucker, 1998).

The above definitions focus only on spoken discourse markers. They are generally thought to be features of oral language, however, Redeker (1991), contrary to his own definition, points out, that there is no reason to believe that discourse markers do not exist in the written language since written discourse also needs to be organized and uses many DMs, especially those originating from the word class, conjunctions. Fraser (1999) defines DMs as follows:

> … a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases. With certain exceptions, they signal a relationship between the interpretations of the segment they introduce, S2. and the prior segment, SI. They have a core meaning, which is procedural, not conceptual, and their more specific interpretation is "negotiated" by the context, both linguistic and conceptual. (p.931)

Fraser's (1999) definition first of all does not limit DMs to either spoken or written language. It also agrees that DMs may be derived from more than one-word classes. It describes the function of DMs as linking the interpretations of different segments, i.e. it is not linking the text of the segments, but rather the concepts or ideas contained within them. Blakemore (1987) makes the same argument by stating that DMs work to link the propositional content of different discourse segments. Fraser also believes that DMs do not carry propositional meaning, something conceded by Hölker (1991, as cited in Jucker, 1998). Some (e.g. Redeker, 1991) do suggest that they can carry propositional meaning, however, a vast number of researchers hold the belief that such is not the case. Finally, Fraser's definition aptly describes the process of ascribing meaning to DMs as dependent on the linguistic and conceptual context of the situation, which shows how the same DM can vary in meaning depending on the situation.

As seen in the definitions, there are various perspectives regarding DMs in the research cannon. As such, there are also many different theoretical approaches that have been taken by researchers to describe DMs. These are discussed below:

2.3 Theoretical Approaches to Discourse Markers

Exploring the different approaches that researchers have taken to describe the role of DMs in language provides us a deeper and more nuanced understanding of

their properties and role in language use. The following are the four major approaches to DMs:

2.3.1 The Coherence Model

Deborah Schiffrin (1987) was the first linguist to properly explore and highlight the role of DMs in discourse. Her understanding of DMs gave birth to the coherence model of DMs, which suggests that DMs serve to organize the discourse in sequential order. DMs are 'contextual coordinates' that help in locating the context of an utterance on both local and global levels. It can be said that these coordinates perform their function in two different ways. First, they help in locating the utterance in terms of the sequence. Secondly, they help in locating the utterances in the five planes delineated by Schiffrin (1987). Each plane has its own coherence. These planes have been discussed below:

- Exchange structure – enables turn-taking in a conversation. It includes conditionally relevant parts that signal adjacency, such as questions and answers and greetings.
- Action Structure – locates the utterance in terms of its sequence within the discourse. It indicates the utterances that follow and precede the current utterance and creates coherence by creating a linear sequential link between the utterances.
- Ideational Structure – the ideational structure is semantic in nature. It creates cohesion by interpreting new information in terms of information in the preceding text. It creates relationships between ideas as well as organizes the discourse in terms of topics and subtopics. It also indicates the functional relationships between ideas.
- Participation Framework – reflects the various levels of identities carried by the talk in a conversation. It reflects the relations between the speakers and the listeners. Furthermore, it also reflects the relations between them and the utterances.
- Information State – it is a cognitive element of discourse, which is utilized by the speakers and the listeners to organize and manage the knowledge and meta-knowledge within the discourse. It relates to the current and evolving knowledge of both parties in a discourse.

According to Schiffrin (1987), DMs create coherence by locating the utterance in one or more of these planes. In addition, they ascribe each utterance to a speaker and/or a listener as well as to its context within the discourse.

This model is a good starting point for the analysis of discourse markers and carries many weaknesses. Redeker (1991) provides a critique of the model, pointing out many flaws, including that the planes of talk are shrouded in relative haziness and need more clarity. For instance, she argues that the ideational structure is unclear and vague with some elements repeating across various planes. She gives the example that ideational structure seems to encapsulate functional relations at first; however, they are later introduced as a part of action structure. Redeker (1991) introduces her own model based on Schiffrin's model. She eliminates two planes: participation framework and information structure, arguing that they are not directly related to coherence. They are embedded within the other planes and thus should be integrated in them. Her model has the following three components:

- Ideational Structure – similar to Schiffrin's ideational structure, it describes the relation between two discourse units based on the speaker's commitment to that relation. This means that the units are related if the speaker believes that they have some kind of sequential relation in terms of time, cause, consequence etc.
- Rhetorical Structure – similar to Schiffrin's action structure, two discourse units are considered to have a rhetorical relation if the strongest relation between them is found between their illocutionary intentions
- Sequential Structure –two discourse units are considered to have a sequential relation, if the relation between them can be described as a transition between topics or a transition towards commentary, correction or paraphrase etc.

Redeker (1991) says that any utterance has elements of all three planes at all times, however, for each utterance in the discourse, one of the planes usually dominates. She also provides her own definition of DMs, as discussed above. The major difference between her and Schiffrin's definitions is that the concept of DMs does not necessarily make them sequentially dependent and does not limit them to bracketing units of discourse. Fung and Carter (2007) also criticize Schiffrin's model for her lack of attention to the interpersonal functions of DMs. They give the example

of the DM “I think” and argue it can act as a hedge and therefore shows a speaker’s intention to not appear forceful, an interpersonal function. Both of these coherence models provide a strong footing to explore DMs, however, they both focus too much on textual coherence and lack focus on the interpersonal functions of DMs. Redeker’s (1991) planes of talk also need further clarity and leave understanding to be desired.

2.3.2 Fraser’s Model

Fraser in his various papers (Fraser, 1993, 1996a, 1999) provides a grammatical pragmatic perspective on discourse markers. His approach to DMs views them in a wider context of the discourse, rather than limiting their function to the textual structure. In his theory, consistent with Schiffrin (1987), he believes that DMs do have a core meaning, which is negotiated by the context it resides in. However, in contrast to her, he does not include non-linguistic expressions in DMs and limits them strictly to linguistic tokens. Fraser proposes that instead of only forming links between consecutive utterances, DMs signal the relationship between the intentions of consecutive utterances. They have a core meaning, which is procedural and not conceptual. Procedural means that their meaning only works to show how an utterance should be interpreted in relation to the preceding utterance. A DM imposes certain interpretations on an utterance based on the preceding utterance and forces it to be interpreted in the context of the intention of the previous utterance.

Aside from excluding non-linguistic expressions, Fraser goes further to exclude some linguistic forms from the category of DMs that Schiffrin (1987) had included. Within these, he especially discusses adverbials, such as ‘frankly’, which he believes do not signal an adjacency pair relationship and rather start a comment, which is a completely new message. This is in contrast to what we have previously seen in Redeker’s (1991) sequential structure, which considers such elements to be a part of DMs. He also excludes pause markers and interjections for similar reasons. These elements go against his belief that DMs only have procedural meaning; therefore, anything bringing new meaning cannot be included in DMs.

Fraser (1996) delineates four classes of discourse markers, namely topical change markers; contrastive markers; elaborative markers; and inferential markers.

- Topical change markers –serve to transition the attention to a different topic or a temporally related event. Some of the examples he includes are the following, “back to my original point, before I forget, by the way, incidentally, just to update you” (Fraser, 1996b, p. 187).
- Contrastive markers –signal that the relationship between two utterances is contrastive, i.e. the following utterance is either in denial of or opposite to the previous. His examples include “all the same, anyway, but, contrariwise, conversely, despite (this/that), even so, however” (Fraser, 1996b, p. 187).
- Elaborative markers – are used to improve upon or explain the previous utterance. Fraser (1996b) provides the following examples: “also, alternatively, analogously, and, besides, better, by the same token” (p. 188).
- Inferential markers – signal that an utterance draws a conclusion from the previous discourse. They usually provide casual relationships. Fraser (1996b) provides the following examples: “accordingly, after all, all things considered, as a consequence, as a logical conclusion, as a result” (p. 188).

Fraser’s model does not take into account parts of the language that do not necessarily segment the discourse but play an important part in setting the dynamic between a speaker and a listener (Jones, n.d.). These include response tokens, e.g. “I see” or “Right” which can arguably add coherence. In addition, similar to Schiffrin (1987) and Redeker (1991), Fraser also fails to properly take into account the interpersonal functions of DMs.

2.3.3 Blakemore

Another theoretical perspective about DMs comes from Blakemore (1987), who works within the relevance theory framework of Sperber and Wilson (1986). In her model, the function of discourse connectives, as she calls them, is to show the dependency relation of one discourse segment on another. Coherence-based theories of Schiffrin (1987) and Redeker (1991) consider the primary function of DMs to bring coherence to discourse, however, Blakemore suggests that it is only one of their functions and their primary function is to allow the formation of inferences based on the interpretation of the utterances. She considers coherence to only be a secondary function of discourse markers and suggests that it does not have any cognitive reality.

The listeners, upon listening, use the DMs to determine if it is relevant to the context. In Sperber and Wilson's (1986) framework, any hearer is entitled to receive information that is relevant to the current context and they will not have to work hard to place the utterance in that context. The function of DMs is to make both of these functions smoother, therefore aiding in creating relevance. She agrees with Fraser in suggesting that DMs do not carry any representational meaning. Their meaning is only procedural; DMs do not carry any meaning but provide instructions to the hearer on how to interpret the utterances to make them relevant. Her concept is a reinterpretation of Gricean notion of conventional implicature.

According to her, DMs are constraints on the contexts of the utterances. They indicate how an utterance is relevant to its context. DMs indicate relevance in four ways:

- By allowing the derivation of implicature from the context
- By allowing the speaker to provide more evidence to strengthen what they said
- By allowing the speaker to contradict what they said
- By allowing the speaker to assign a role to the utterance (Blakemore, 1992).

2.3.4 Functional Perspective

Most theoretical perspectives until now have focused on grammatical coherence and pragmatic elements of discourse. They have mostly ignored the functional aspects of DMs. Some key analyses in terms of a functional perspective have come from Aijmer (1996, 2002) and Fung and Carter (2007), with both having done corpus-based studies of DMs. Aijmer's (1996) work has mostly focused on speech act adverbials.

According to Aijmer (2002), DMs have a strong potential for indexicality, i.e. the ability to point. They derive their meaning from the context and then use it to connect the discourse. She shares Fraser's belief that DMs only have procedural meaning. They are assigned a core meaning when they come into a context and derive their meaning from there. DMs work to place utterances into a conversation. They allow the listeners to interpret the utterance according to the context by serving as restrictive elements. She agrees with Schiffrin's idea of describing DMs on different

planes, however, she also points out that Schiffrin's planes of talk lack clarity. She also proposes that DMs should be analyzed on textual and interpersonal levels.

Aijmer (1996) proposes two functional classes of DMs for analysis: global and local discourse markers. Local DMs are found in the flow of the conversation. They provide information about the relationships between adjacent utterances. Global DMs create global coherence in the discourse and showcase the evolution of the discourse.

She also provides a classification of DMs in terms of their deictic orientation (Aijmer, 1996):

- speaker-oriented (e. g. I mean, I think, in my opinion) – showing a speaker's own attitude towards the discourse
- hearer-oriented (e. g. now you come to mention it) – showing some kind of reaction to the hearer
- speaker and hearer-oriented (e. g. let's put it, let's face it) – involve both the speaker and the hearer into the conversational stakes
- third person-oriented (e. g. as far as X [third person] is concerned) – used to point to a third person

Aijmer's description of DMs has a lot of strengths. First of all, she gives due importance to the interpersonal functions of DMs in discourse, by making them one of the two planes of analysis. Her method of analysis also holds greater clarity than earlier models.

Fung and Carter (2007) also provide a similar kind of analysis to Aijmer. In their own corpus-based analysis, they identify four different classes of DMs:

- Interpersonal – indicate shared meaning (*e.g., See, you see, you know, listen*), signal attitude *(e.g., Well, really, I think, obviously, absolutely, basically),* and show responses/reactions *(e.g., OK/okay, oh, right/alright, yeah, yes,).*
- Referential – provide cause *(e.g., Because)*, contrast *(e.g., But, however)*; perform coordination *(e.g., And)* and disjunction *(e.g., Or)*; show consequence *(e.g., So)*, digression *(e.g., Anyway)* and comparison *(e.g., Similarly)*.
- Structural – appear in opening and closing of topics *(e.g., Now, OK/ okay, right/alright)*; indicate sequence *(e.g., First, firstly, second)* and topic shifts

(e.g., So, now, well); summarize opinions *(e.g., So)*; and indicate continuation of topics *(e.g., and, cos, so)*.

- Cognitive – denotes thinking process *(e.g., Well, I think, I see, and)*; allows for self-correction and reformulation *(e.g., I mean, that is, in other words)*, and elaboration *(e.g.,* Like, I mean); show hesitation (e.g., Well, sort of); and show an assessment of the listener's knowledge (e.g., You know).

2.4 Characteristics of Discourse Markers

From the above discussion, we can see that it is not possible to agree upon a single definition of DMs due to different approaches to their studies adopted by various linguists. However, for the purposes of the study, we need definitive criteria to choose DMs. The way forward may be to work with some characteristics of DMs that seem to be common among different approaches. Fung and Carter (2007) provide the following five criteria for DMs in their article:

2.4.1 Position

The majority of DMs appear at the start of an utterance or at turns in a conversation. With the initial position, they mark boundaries between topics and thoughts as well as signal reactions and orientation. However, according to Fung and Carter (2007) it is not necessary for all DMs to appear at the start and can appear at other positions. Examples include the following:

a) **As far as I am concerned**, there is no need to pay any attention to this racket.
b) I am going to the market. **However**, I will not be able to shop for art supplies.
c) He is not going to like the color **in my opinion**.

2.4.2 Prosody

DMs are independent of the utterance in terms of their prosody. Their prosodic features, such as intonation, pauses, etc., are usually different from the rest of the sentence. They are separate tone units, a feature that aids in the achievement of their function in organizing the discourse.

2.4.3 Multigrammaticality

DMs do not form one single grammatical group and emerge from several different linguistic classes. Fung and Carter (2007) provide the following examples of the categories which contribute to the class of DMs:

> coordinate conjunctions (e.g. and, but, or); subordinate conjunctions (e.g. since, because, so); prepositional phrases (e.g. as a consequence, in particular, by the way, at the end of the day); adverbs (e.g. now, actually, anyway, obviously, really, certainly, absolutely); minor clauses (e.g. you see, I mean, you know); response words (e.g. yeah, yes, no); interjections (e.g. oh, ah, well); meta-expressions (e.g. this is the point, what I mean is, that is to say, in other words) (p. 4).

It must be noted that while DMs do emerge from these categories, not every item from these categories will necessarily be a DM.

2.4.4 Indexicality

Indexicality means the ability to point. DMs are able to reference or provide directions to the context of the utterance and therefore allow linking of the discourse units. These DMs can range from being conceptually empty (e.g., ok, oh) to being conceptually rich (e.g., I think, frankly). Fung and Carter (2007) further postulate that even conceptually rich items go through gramaticalization to settle into their indexical function, which could either be textual or interpersonal.

2.4.5 Optionality

Finally, the last characteristic described by Fung and Carter (2007) is optionality. Since DMs do not contain any propositional meaning, their absence does not affect the meaning of the text. While their absence will definitely affect how a listener or reader perceives the text in context of the rest of the discourse, they can still be understood standalone.

With the above discussion, a further look into the kinds of DMs may provide a clearer picture of their nature and group them into categories that are easier to work with.

2.5 Types of Discourse Markers

Just as the work on other aspects of DMs is varied in each researcher's work, DMs have been grouped into different categories by different researchers. The following classification relies on the division by Fung and Carter (2007), as it sufficiently covers most of the other established classifications. They have provided the following four categories or types of DMs:

2.5.1 Interpersonal

Interpersonal DMs have to do with the social and affective elements of speech or text. They are used to perform two functions. The first is to mark different responses, such as agreement, acknowledgment and confirmation. These include markers such as 'ok/okay', 'I see', 'great' etc. The second function is to signal attitudes and stances regarding the subject. Maschler and Schiffrin (2015) describe them as elements that show the relation of the speaker with the listener as well as the text. These include markers such as 'I think', 'to be frank', 'actually', 'basically', 'go on', 'woah' etc.

2.5.2 Referential

Referential DMs mark the relationships between verbal activities in consecutive utterances and usually comprise conjunctions, such as 'because', 'however' etc. Maschler and Schiffrin (2015) describe them as 'textual' discourse markers, which indicate the relationship between subsequent utterances. These relationships may include cause, coordination, contrast, disjunction, digression or comparison. Examples include "now". "so", "therefore", "because" etc.

2.5.3 Structural

Structural DMs perform the functions of facilitating the transition between topics and turn-taking. According to Fung and Carter (2007), they indicate the beginnings and endings of topics with markers, such as 'let me start by' (beginning) and 'to conclude' (ending), demonstrate sequential relations with markers such as 'first of all' and signal shifts in topics with markers such as 'what about'.

2.5.4 Cognitive

Fung and Carter (2007) propose another category of DMs, called cognitive DMs, which are used to show the cognitive state of the speaker. They perform functions such as giving the speaker time to think (e.g., 'I see'), allowing them to reformulate (e.g., 'in other words'), elaborate (e.g., 'I mean'), and assess the listener's knowledge (e.g., 'you know'), reacting to information (e.g. 'oh').

For the purposes of our current study, the following section explores the landscape of teaching DMs and discusses the merits of well-known studies.

2.6 Previous Attempts at Teaching Discourse Markers

In terms of broader approaches, there are two major ways that linguists have experimented with in teaching DMs: inductive and deductive. The deductive approach represents the more traditional way of teaching, reflected in models such as the Presentation-Practice-Production (PPP) model. This approach is characterized by the explicit provision of explanations for the structures being taught (Shaffer, 1989). The inductive approach, on the other hand, finds its way into methods such as the Task-Based Language Teaching (TBLT) method, Problem-Based Learning, etc. This approach advocates for a method where the students are made to focus on the structures being learned and try to figure out the patterns found to reach a conclusion about the underlying rules (Shaffer, 1989).

Rahimi and Riasati (2012) conducted research on the explicit teaching of DMs with 40 Iranian English language learners. The research was focused on teaching spoken discourse markers and included research subjects between the ages of 23 to 30 years. The control group received no teaching of DMs, while the experimental group was treated with explicit teaching. Pre and post-tests were run in the form of spoken interviews where the students were made to talk on certain subjects. The results indicate a significant improvement in the use of DMs in the experimental group. This study also claims that implicit teaching of DMs did not yield any results. However, looking at the method of treatment of the control group, it appears that the control group was merely taught the English language normally and no strategies for the implicit teaching of DMs were employed to say with certainty that implicit teaching does not work. This, however, does point to the fact that the methods of implicit or inductive teaching need to be more clearly defined to ensure that the learning goals

are accomplished. This study, moreover, also provides evidence that the teaching of DMs does indeed benefit learners.

Similar to the above study, Asl and Moradinejad (2016) conducted a study on 41 university students by dividing them into an experimental and a control group with 21 and 20 participants respectively. This study was also focused on spoken DMs and treated the experimental group with explicit teaching while the control group received no explicit teaching of DMs. This study reflects the results as well as the flaws of the study discussed above, where it found that explicit learning led to improvement in learners' ability to use DMs. However, no delayed post-test was administered leaving, therefore no data is available regarding the long-term retention of the concepts. It can be argued that these studies do make a case for teaching DMs, as it seems to benefit learners' use of DMs, however, long-term retention is an area that leaves wanting for more research.

Another study on Iranian EFL learners by Khandaghi Khameneh and Fakhraee Faruji (2020) takes 60 learners and treats them to the explicit teaching of DMs. The learners in this study were placed at the intermediate level and were younger than the average participants in the study conducted by Rahimi and Riasati (2012), where the age range was 23 to 30 years old. The research instruments were speaking pre and post-tests in the form of interviews. The experimental group was treated with three-step lessons where they had to listen to conversations and take note of the DMs. They were, then, provided explanations about the functions of the DMs. The control group only received a standard English language lesson. The results of the study showed no significant difference between the control and experimental groups. These results are inconsistent with the study we have reviewed above. One important factor that might be the cause of the difference here is the difference in age as well as the proficiency of the students. The participants in Rahimi and Riasati's (2012) study were older and were participating in a program for IELTS, in comparison to the participants in this study, who are students with intermediate English proficiency. It could be hypothesized that the teaching of DMs works better for learners who have a better base level of proficiency.

Sadeghi and Kargar (2014) studied the effect of explicit teaching of DMs on 40 learners ranging from age 12 to 17, with 20 of them each in the experimental and

control groups. Unlike the studies above, this study focused on written discourse markers. The learners were exposed to 15 sessions. They were expected to write 150 word essays in pre and post-tests. This study also found a significant difference between the experimental and control group with the experimental group showing a significantly higher level of progress. The study also noted that younger pre-intermediate learners showed lower levels of progress in comparison with older intermediate learners, which is consistent with the previous results where we hypothesized that older learners respond better to the teaching of DMs.

Jones and Carter (2014) conducted a study in which they compared learners exposed to explicit and implicit methods of teaching. It comprised 36 participants divided across three groups: two experimental groups, one being taught with the Presentation-Production-Practice (PPP) method and the other being taught with the Illustration – Interaction – Induction (III) method while the third group, the control group, was not taught anything at all. The control group was only exposed to an English-speaking environment, as the participants were Chinese learners who had been in England for three weeks on average before the start of the study. All three of the groups received lessons over five days and were tested before and after them. The results of the study show that the learners exposed to the PPP method had the greatest amount of improvement, which was carried over to their delayed post-test results, where the group again scored the highest. However, a deeper dive reveals that the results may not be most reliable. The biggest concern in relying on the results of the study is that the PPP group scored 19 on the pre-test and scored 39 on the post test, while the III group scored 3 on the pre-test and scored 15 on the post-test. If only seen in terms of percentage increase, then the III group has shown significantly better improvement. This unreliability can be further seen by observing the behavior in the control group, which scored 10 in the pre-test, 3 in the post-test and 11 in the delayed post-test, demonstrating that such a small sample size may not be statistically reliable. This leaves room for more studies to be conducted to ensure more reliable results.

Moghaddasi, Bavali and Behjat (2020) compared the impact of teaching DMs on the speaking skills of learners by either implicit or explicit means. To this end, they divided 90 learners into two equal experimental groups, each taught with either the implicit method or the explicit method for eight sessions. There was no control group for the study. Data was collected using interviews before and after the study

treatment. The experimental group with explicit teaching was taught using a method, which involved explaining the use and functions of the DMs to them. On the other hand, the second experimental group did not receive any explicit teaching about DMs and were implicitly taught using only corrective feedback on the DMs in their speech. The study found that while the improvement for both groups was significant, the group taught using the explicit method presented better scores in the post-test as compared to the implicit method. The results of this study make a case for the explicit method to be better; however, there are a few areas of concern. The first issue is the absence of a control group for the study. Both of the groups are experimental and there is no independent group that could serve as control, therefore leaving the results of this study less reliable. Secondly, the implicit method of teaching here only comprised of providing corrective feedback on the learners' use of language. While the explicit method raises the consciousness of the learners regarding the use of the DMs, the implicit method used in this study seems inadequate for the learners to notice the patterns of their use. Moreover, merely practicing language in a class may not be enough for learners to start acquiring the use of language, unless they are provided a large number of authentic examples, which put the target word into focus. While this research makes a case for explicit learning, it allows us to conclude that the methods for implicit learning need to be reevaluated to make sure that it actively serves to raise consciousness for learners.

Alraddadi (2016) also explored the difference in the teaching of DMs through inductive and deductive methods. The subjects of his study were 41 female EFL learners. It divided the students into two groups: one treated with TBLT and the other with PPP. The study found that in the post-test, both methods were equal in lifting the use of DMs in learner language. However, the key difference emerges in the delayed post-tests, where TBLT takes a notable edge on PPP in terms of retention and long-term use of DMs.

Contrary to the results of the above study, the study conducted by Kapranov (2018) in Ukraine with adult EFL learners in a standalone EFL course had different results. The research involved teaching learners two hours every week for five months. One of the groups in this research was taught by allowing them to explore DMs in the course texts and other course-related activities. The control group, on the other hand, received explicit instruction on DMs. The study found that the learners in

the experimental group did not show any improvement, compared to the learners in the control group which did show improvement. The results of this study are in contradiction to what the above study by Alraddadi (2016) shows. We have also seen this in the first study we discussed, which did not find any significant improvement in learners who are taught with the implicit method. The above review of literature provides two possibilities about implicit teaching of DMs, with the first being that implicit teaching of DMs is ineffective and the second being that the current methods of teaching using the inductive methods are unrefined and need more exploration to refine the methodology. The second possibility seems more plausible here as it would be intellectually dishonest to rule out implicit teaching when there are studies that have reported positive results. This means that the implicit method needs more research to agree upon a method that could be the most effective.

It is also essential to add further context to the teaching of DMs by looking at the studies conducted in Pakistan, as it is the target geographical area for the current study.

2.6.1 Teaching DMs in Pakistan

There has been little research conducted on testing different methodologies in teaching DMs. Most of the studies in Pakistan are cross-sectional studies which show how Pakistani English language users from different contexts use DMs and compare them to other demographics (e.g. Jabeen et al., 2011; Malik et al., 2021; Nawaz et al., 2021; Noor, 2021; Shafqat et al., 2020; Sultan et al., 2021). However, the work on actually teaching DMs is very low. There was only one prominent study found on the subject conducted by Yasmin et al. (2021). Their study consisted of 40 university students divided into two groups: one experimental that was explicitly taught DMs and one control group that received no treatment. The study measured learner proficiency in the pre-tests and the post-tests using an IELTS interactive test. The participants of the two groups were equal in the pre-test, however, learners undergoing explicit instruction showed higher scores in the post-test compared to the participants of the control group. The study concluded with the interpretation that explicit instruction is indeed beneficial for learners; however, the study did not provide any activities for the learners to engage in implicit learning. Moreover, the study did not conduct any delayed post-test for inquiry into long term retention.

2.7 Data-Driven Learning

Various studies have proposed a DDL-based method of teaching as an excellent way for inductive/implicit teaching (Cobb, 1997; Johns, 2002, 1991) due to its exploratory nature. It is based on the idea that learners should be exposed to authentic language and should be allowed to figure out the rules by themselves like detectives. It puts the learner into the driving seat of language learning and urges them to work on the data like researchers (Johns, 2002). This is mostly achieved through concordances which provide context for specific keywords. Before going on to explore the approach itself, let us first dive into its history.

2.7.1 History of the Data-Driven Learning Approach

Alex Boulton (2011) traces the development and the history of the Data-Driven Learning approach through the past decades. According to him, there are conflicting reports about the earliest use of corpus in language teaching. While the invention of the term "Data-Driven Learning" is attributed to Tim Johns, who has been the pioneer of this approach, McEnery and Wilson (1997) report the first usage of corpora in language teaching to be by Peter Roe, a researcher at Aston University in 1969. Johns himself in his own 1986 study cites Ahmad et al. (1985) as the first group of researchers who used corpora for such purposes. Another earlier publication that seems to deal with the application of corpus in language teaching comes from Sandra McKay (1980), who experimented by providing extracts of texts to her learners. Nevertheless, Johns has been credited as the pioneer of the field and responsible for setting out a majority of the concepts related to the field. He discussed these concepts in a variety of his publication including “nine articles (T. Johns, 1986, 1991, 1993, 1998, 2002; T. F. Johns, 1991), a co-edited collection (Johns & King 1991), a number of online collections, and a short discussion in a co-edited book (Higgins & Johns 1984)" (Boulton, 2011, p5).

Boulton (2011), a key figure in DDL research, goes on to discuss the evolution of DDL through Johns's own works as well as other researchers. Johns himself changed his opinion about DDL a few times. Initially, he believed it to be a set of techniques, which could be integrated with other older methods of teaching. However, he later changed his stance and started calling it an approach. DDL also took several different names in different articles presented by Johns including classroom

concordancing, corpus-driven approach, and corpus-based approach. Regardless of what name it took, the idea behind it was the same every time: the use of texts as data for learners to exploit.

Boulton (2011) also discusses the variety of methods and features under the umbrella of a data-driven approach. He points out that many studies have used the name of Data-Driven Learning; however, they show immense variety in how they deal with it. It ranges from completely hands-on experiences to printed and extracted forms of text. They have used single text corpora to texts with hundreds and thousands of different texts. In short, the methodology has not been consistent and it could be said that the only common feature among most of the studies was the common use of a corpus for language teaching.

Before going on to investigate how DDL works, it must be established why DDL could be a viable option for language teaching. The below section discusses the merits of DDL which bring it into consideration for a language teaching approach.

2.7.2 Why Data-Driven Learning

English language teaching and learning has found itself in the application of many different teaching strategies and methods. Linguists have painstakingly described the grammar of the language and have devised methods to teach these grammatical rules to learners. A majority of these methods has been historically based on explicit teaching of rules and grammar, so that the learner can form a complete abstract knowledge of the language. Data-driven learning, however, takes an entirely different view of language teaching and disposes off rule teaching in favor of pattern searching and there are good reasons for it. Boulton (2007) points out that the concept of rule teaching is itself artificial. Humans are inclined to learn things in the environment by noticing patterns within the universe, whether they are related to animate beings or inanimate things. Rules were devised by humans themselves through the process of noticing patterns, and were believed to allow other learners to skip the process of figuring out these rules by themselves. However, such a notion might be flawed, since traditional teaching methods have not yet been successful in achieving a high rate of proficient language use. Many learners can build an abstract knowledge of the grammar, but they are unable to apply it in their own language use.

Human brains have evolved to notice patterns and are adept at it naturally. Boulton (2007) cites the example of the Wason card selection task, where those presented with the task fare better if they are presented with the cards and asked to find the pattern of choosing the correct cards, rather than explaining the rules. However, Boulton (2007) also points out that leaving the learners with the target material cannot be enough either. He argues for a teaching methodology where the teacher plays the role of a guide or research organizer, who could lead the learners through the process of finding out grammar rules. Moreover, the learners may not always be able to voice or formulate the rules like experts, however, it does not mean that they do not understand the rules and a better reflection of their ability is their actual use of the words. Another notable thing here is the finding that learners do not always acquire language rules in the same sequence as they are taught, because there seems to some innate system in places which guides the order of grammar or lexical items to be acquired first (Ellis, 1993), which raises more questions about the effectiveness of rule teaching and leaves room for us to explore whether an implicit method of learning will be able to come more naturally to the learners.

2.7.3 How Data-Driven Learning Works

DDL is a corpus-based approach, which means that there are two essentials that make DDL: the corpus and the corpus tool. Both, the choice of the corpora as well as the corpus tool depend on the target use. For the purposes of DDL, a corpus tool with a corcondancer usually suffices, because DDL mainly relies on the KWIC (Keywords in Context) function of the corpus tools, which provides all the instances of a keyword occurring in the corpus. The choice of a corpus also holds an important place in DDL due to its own focus on the authenticity of the text. Granger (2010) discusses that while the corpora always consist of naturally occurring texts, they may not be always be authentic in text reception (Widdowson, 2000 cited in Granger, 2010). This means that the learners may not always be able to relate to the text in a corpus that may not have been specifically designed for their use, which could hamper their learning. She offers two kinds of corpora as a solution for this problem: the pedagogic corpus and the local learner corpus. A pedagogic corpus includes text from the textbooks of the learners. Textbooks are usually designed by the relevant experts of each country and are therefore culture and level appropriate. A particular advantage to such texts may also be the fact that students might have already studied these texts

in classrooms, therefore they have better context to understand them. On the other hand, a local learner corpus comprises of text produced by other non-native local learners of the language. Such a corpus reflects the interlanguage features that are found within the language of the learners themselves. This provides a chance to the learners to analyze language produced by learners with the same mother tongue as them.

Another important aspect of the use of corpora is annotation. Corpora can be annotated for various features, such as Part-of-Speech (POS) tagging, phonetic tagging, semantic tagging, error tagging, etc (Granger, 2010). Each kind of tagging is used for different educational goals. According to Granger (2010), tagging can be problematic too, because it always reflects a theoretical perspective of the language, which the teacher might not always agree with. Tagging can also be tedious and is difficult to be completed individually by every teacher. The internet has now enabled corpora to be shared across the world, which also come annotated for several different characteristics. This has paved the way significantly for the use of DDL in the classroom.

According to Granger (2010), the decisions about what to present and what to do in a DDL classroom are dictated by different factors, including the “learning context”, “the level of learners”, and “the topic investigated” (p. 4-5). These factors also dictate the choice of tool that will be suitable for the respective activity. Most DDL activities rely on the concordances that provide KWICs. KWICs provide instances of the same word in a whole corpus. While the lack of further context than a sentence can be off-putting to some learners, it greatly helps them in noticing patterns, since they can see how a word behaves in different situations. Granger (2010) also emphasizes the need of editing the concordances to make them more suitable for student use. Hadley (2002) demonstrates the use of KWICs for learners to find out the correct use of a language item to point out that sometimes the use of a word might be grammatical, but it can still be wrong as it might not collocate with the words it is used and therefore, sound unnatural to the listener. Presenting a large number of examples of the same word to the learners allows them to regulate their language and ultimately generate appropriate language by themselves.

The above section discusses the theoretical work on the DDL approach, however, it is crucial that the application of DDL as a language teaching approach is looked into, as it directly provides information about the applicability of the approach in actual classrooms.

2.7.4 Previous Work on Data-Driven Learning

Although the number of empirical studies of Data-Driven Learning (DDL) is still less than a lot of other approaches, there still have been many efforts to study its effects in the real world. A study by Ming Huei Lin (2021) studied the effects of teaching grammatical features using DDL in control group and treatment group settings. The treatment group comprised 95 first-year undergraduate learners while the control group consisted of 84 learners. Following the concepts of discovery learning and noticing hypothesis, the researcher delivered lectures to the learners about grammar items and using pre-post-tests, found significant improvement in grammar performance in the treatment group. While this indeed does show that DDL can be used to teach grammar items, it does not provide any information about its efficacy in comparison with traditional forms of language teaching, because the researcher did not teach the control group anything at all.

A similar study by Fang et al. (2021) tested 22 senior secondary school learners in China, who wanted to appear for the International English Language Testing System (IELTS) test in the future. The research found that in just three corpus-based training sessions, learner proficiency significantly increased, as measured by the pre and post-tests. The training sessions were hands-on. The researchers performed an error analysis on the pre-tests of the learners and guided them to explore the collocation errors found in those tests using corpus tools. These activities resulted in significant improvement in the writing quality. This study, however, also lacks a control group to measure the benefits of using DDL in the classroom over other approaches.

Vasiliki Papaioannou (2018), for her doctoral thesis, did a comparative study of teaching writing skills to Greek learners using both DDL and the traditional course-book approach. She used a blended form of DDL, in which she combined the hard and the soft versions of DDL, the former of which only includes hands-on experience with corpus tools and software, while the latter allows printing out concordances to be

distributed to learners. Using pre-tests, post-tests and delayed post-tests on control and experimental groups, she found noticeable improvements in the group taught using the DDL approach in both the post-tests and the delayed post-tests. Another significant finding in her work was that learners with lower levels of beginning proficiency benefited more from DDL. This finding is especially interesting, since DDL has been long considered as an approach that is not suitable for learners at lower levels (e.g. Johns, 1986).

The above research clearly carves out a space for a DDL-based study in the Pakistani context. Data-driven learning covers a lot of the flaws of the earlier approaches. It can be used to teach all four language skills. Since it is an inductive technique, it stands to reason that it provides a longer lasting and more comprehensive understanding of grammatical concepts. In addition to that, it covers the weaknesses of earlier inductive approaches, which used fabricated, artificial examples of language to teach learners. As we discussed earlier, the communicative approach came under criticism that situations covered in the classrooms were not reflective of actual real-life language use, which is resolved by using authentic examples from the corpus. Similarly, DDL does not discount the role of the teacher and puts explicit focus on the presence of the teacher in the class as the facilitator, which would guide the learners through discovery, thus allowing better use of time and more directed learning in the class. This discussion, however, begs the question of its applicability in the Pakistani context. The following section, therefore, discusses the history of the approach in Pakistan.

2.7.5 Data-Driven Learning in Pakistan

While DDL has its roots in the 1980s due to the work of Johns, it still has taken a considerable time to find its way into empirical research and application in classrooms due to its unorthodox nature. Corpora have been used in many studies situated in Pakistan; however, these were cross-sectional studies focusing on describing the current state of things with very few studies on using corpus for teaching language. The study conducted by Shah (2021) was situated in GC University Faisalabad where 100 participants took part. They were divided into an experimental and a control group, with 50 members each. The experimental group was made to undergo an instructional treatment with computer assisted DDL to study its effects on lexical collocations. The analysis of results revealed that the

experimental group had higher scores than the control group and provided basis for believing that computer assisted DDL could be beneficial for learners at university levels.

An article published by Zahra and Abbas (2018) is a qualitative study reflecting on the potential implications of corpus based approach for language teaching in Pakistan. The study explores how the Michigan Corpus of Academic Spoken English (MICASE) can be used in Pakistan to teach different lexical items. It does not directly deal with its implementation by experimenting in class and therefore, only serves as a theoretical look into the possibilities of teaching.

2.7.6 Reservations about Data-Driven Learning

Data-driven learning has always faced reluctance in adoption by both linguists as well as educators. There are many issues behind this. Gilquin and Granger (2010) flesh out four major aspects that might be problematic: "the logistics, the teacher's point of view, the learner's point of view and the content of DDL" (p. 366). Logistical issues include issues like arranging a corpus, concordance software, computers and other related equipment. The software as well as hardware related to DDL is not only expensive but also requires extra training and expertise to use well. Moreover, software and hardware can be expensive, leaving a lot of schools and universities unable to afford them. While the soft version of DDL requires fewer resources, it still requires worksheets as well as corpus material to be printed, which could also add up to the expenses. The teachers also need to spend time making worksheets as well as to prepare language samples. The logistical aspect also includes concerns about time. Preparation of material as well as acquiring training is taxing on time. Moreover, younger and less experienced learners may need time to familiarize themselves with the use of computers. Corpus software are not always interactive and user-friendly, therefore, even advanced learners could take a fair bit of time to understand it. Another notable thing of concern to be paid attention to the time spent in deriving the rules. Since DDL lets the learners take the role of a researcher, it can be time-consuming, with some learners feeling that it would be faster with the teacher simply explained the concept (Hewings, 2012, cited in *Corpus Linguistics for ELT: Research and Practice - 1st Edition - Ivor*, n.d.).

Coming to the teacher's point of view, as described by Gilquin and Granger (2010), teachers find it hard to adapt to and adopt DDL in their classrooms due to unfamiliarity. For most teachers, it is a unique and new method that requires a new set of skills and extra time set aside to learn them. There may also be a need for training programs during their service. On the flip side, learners may also find it hard to adapt to corpus software, since it needs a rather significant amount of effort to not only learn the software and the computer skills but as well as to act independently as researchers. Learners may not be adept at operating independently, even with the teacher guiding them and may need time getting used to it.

The learner-led nature of the DDL method may also be a cause of concern for teachers since traditional schooling methods are very risk-averse. Handing over the control of learning may seem risky, because it could easily lead to errors, even though learners can eventually correct themselves with the guidance of the teacher. This loss of control may also be perceived as a loss of skill by the teacher, since a DDL classroom looks vastly different than a traditional classroom and the teacher's role is limited as a facilitator.

DDL may also be off-putting from a learner's point of view, since it requires them to acquire a very different set of skills compared to normal schooling. These skills include corpus literacy, i.e. the ability to use corpora and corpus software, as well as the ability to conduct research. DDL carries the risk that a learner might not be able to derive correct rules or even any rules from the corpus. The success of DDL also relies on the learning style of the learner, where independent, intuitive learners may excel but others may not.

Finally, Gilquin and Granger (2010) also point to the content of DDL as an area of problem. The content provided by DDL may not be consistent at all. A streak of strings might only show one type of usage of a language item. Some corpora might not have specific kinds of usages at all. It is also possible that there might not be enough examples of a kind of usage for a learner to be able to recognize it as a separate pattern. Since normal language use also includes some language that is not the standard, it could mislead the learner about normal usage. It is also easy to get swept in the illusion of frequency and consider it the only important measure for the importance of a grammatical item.

While all of these concerns about DDL are legitimate, these should be seen as obstacles to be overcome through experimentation and innovation, rather than reasons for abandoning the method. Logistical issues can be expensive and time-consuming to solve, however, the costs after their initial cost of setup may be reduced significantly. Issues like arranging corpus software, building a corpus, and buying hardware may require hard work during the initial period and may only need slight updation over time as the teacher sees fit. Soft DDL may also provide more relief with places of schooling requiring even fewer resources to set up.

Computer literacy and corpus literacy may also be hurdles in the front, however the penetration of computer technology is more than ever in the current time and it does not present as big of a hurdle as it did even half a decade ago. While training for implementing DDL in classrooms might still be expensive and time-consuming, these skills can be retained by continuous use by teachers. Learners can also become skilled over time if introduced early in their courses. They may also find themselves more interested in DDL due to its novelty and the avenue to explore a new form of learning.

The concern about DDL taking a longer time to introduce concepts than traditional learning needs to be reconsidered. The rationing of time should be reevaluated by keeping in mind the possibility that DDL may lead to better retention and application of concepts, which will provide benefits for further learning in the future. This, however, remains to be explored in depth in further research.

2.8 Summary

The definition of DMs is a contentious matter and many researchers have attempted to describe them in their own ways, an issue that stems from the huge diversity in their origins in the language. However, there are indeed common properties found in the literature that provide a good enough description of their use and properties. Data Driven Learning is a corpus-based method of language teaching which attempts to teach language by exposing learners to examples of natural language and letting them take charge as researchers to figure out the rules. This makes it especially suitable for linguistic items such as DMs, which do not have defined parameters of their origins as was discussed in the above literature.

Besides describing and elaborating the key concepts, the literature provided an insight into the gaps present within the research, especially in research surrounding the Data Driven learning. The research lacks in targeting school level learners, with most of the focus remaining on advanced language learners, especially those in universities. Research in Pakistani context is especially scarce, in respect of both teaching DMs as well as teaching using the DDL approach. This leads us to a path of a new exploration that is reflected in the form of this study.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Overview

This chapter presents an in-depth discussion of the methodologies used in this research. It discusses the theoretical framework that makes the basis of the current study. It, then, delves into the research design, the instruments and procedures, tests as well as the population and the samples used for the research. Furthermore, it also goes into detail about the making and use of corpus in the current study.

3.2 Theoretical Framework

This section discusses the theoretical underpinnings of the current study.

3.2.1 Data Driven Learning Approach

Data Driven Learning (DDL) is a corpus-based approach to language teaching, which believes in giving learners the role of a researcher. It uses the corpus as data for the learners to discover rules from. To broadly define it, it “involves the use of dedicated concordances to explore large language corpora” (Boulton, 2016, p. 268). DDL has the advantage of using examples from real-life language use. Flowerdew (2009) aptly points out that "no dictionary or grammar is able to fully describe the language" (p. 329). Data-driven learning thus bridges this gap by providing authentic linguistic examples. It uses concordance software to bring forth all the examples of a linguistic item and presents it in the form of Keywords in context (KWICs). It relies on the learners' ability to notice patterns of grammar among these examples and form generalization based on them (Johns 1991a). Johns saw the purpose of DDL to expose learners directly to the language. The idea behind these methodologies is consistent with Rutherford's (1987) idea of consciousness raising, which proposes a method of language learning which draws learner attention towards the features of the target language.

3.2.1.1 Consciousness Raising.

Rutherford and Smith (1985) define Consciousness Raising (CR) as "the deliberate attempt to draw the learner’s attention specifically to the formal properties of the target language" (p. 107). It bases language learning on making learners conscious of specific language forms or structures. It encourages the learners to notice

these patterns and internalize them in a manner that would be similar to first language acquisition.

Meunier (2020) argues that DDL's ability to allow access to frequency data and authentic language examples, which are naturally patterned in nature, is underpinned by consciousness-raising, allowing for the learners to discover the rules themselves. The exposure to language examples can draw the learner's attention towards the features and relationships present in the text with the aim that the learner would start inferring the rules. The use of authentic examples makes sure that the inferred rules reflect real-life language use and are applicable in practical use. The involvement of the teacher as the guide as well as the design of the activities also actively directs learner attention towards specific features to enhance noticing and thus, raising consciousness.

3.2.1.2 Types of Data-Driven Learning.

DDL has two major branches in terms of the medium of application: the computer-based hard version and the paper-based soft version. The original form of DDL is the hard version, as it was derived from corpus, which was stored in computers, although Boulton (2010) does find that even Johns, who is credited as the pioneer of the method, frequently used paper-based alternatives in his classes, since the technology was even more scarce at that time.

There are pros and cons of both versions, which have been described by many researchers, albeit mostly in qualitative studies (Boulton, 2010). Hard DDL gives more autonomy to the learners; pushing them more into the role of a researcher. Corpus software provides lots of additional tools for the learners to enrich what they are learning, for example, the ability to obtain way more context for the linguistic item they are studying than paper-based studies. They also have access to more real-life examples of the linguistic item, since only a handful number of KWICs can be printed for soft DDL; however, corpus software can show a larger number of examples. Paper-based DDL takes out a lot of decisions from the hands of the learner, such as choosing which examples to focus on or how much context they need. However, two empirical studies conducted by Boulton (2010) have revealed that paper-based DDL still helps learners discover patterns in languages and productively use them in their language. These, however, were small-scale studies, which warrant

further investigation to support or reject the theoretical stipulations made by various researchers about the advantages and disadvantages of both versions of the DDL method.

There are some other concepts that make up the theoretical basis of Data-Driven Learning. These include the noticing hypothesis (Schmidt, 1990), input enhancement (Smith, 1993) and the involvement load hypothesis (Hulstijn & Laufer, 2001). The noticing hypothesis was presented by Schmidt (1990), who proposed that language learning requires a learner to actively notice and consciously register language input. To convert the input into 'intake', the learner must be focally aware and attentive to the stimuli. Although the noticing hypothesis has come under criticism, its facilitative role in language learning cannot be denied. Schmidt himself amended the hypothesis after the criticism; changing his original position of noticing being essential for language learning to it having a facilitative role (Hulstijn & Schmidt, 1994). Input enhancement is also closely related to the noticing hypothesis and provides techniques that are used to promote noticing (Smith, 1993). Different techniques are used to bring attention to the target features, including highlighting, underlining and coloring.

3.2.2 Noticing Hypothesis and Input Enhancement

The noticing hypothesis was presented by Schmidt (1990), who proposed that language learning requires a learner to actively notice and consciously register language input. To convert the input into 'intake', the learner must be focally aware and attentive to the stimuli. There are three levels of consciousness: perception, noticing and understanding. A learner may perceive some information; however, understanding cannot happen without noticing first. Noticing is the process, through which the mind allocates cognitive resources to stimuli, leading to its understanding and addition to the memory (Robinson, 1995).

The Noticing Hypothesis has been a controversial topic among linguists with many regarding the role of noticing as common sense while others consider it too vague to be given any consideration. The beliefs held by the noticing hypothesis seemingly oppose Krashen’s dual-system hypothesis which postulates that language acquisition happens unconsciously. However, the claims of the noticing hypothesis do not necessarily contradict unconscious learning. The noticing hypothesis claims that it

is noticing that is required for learning and not understanding, which means that implicit learning is still possible, because noticing lies before understanding (Schmidt, 2012). While the hypothesis has come under a lot of criticism, its facilitative role in language learning cannot be denied. This is supported by Schmidt's (2010) meta-analysis of empirical studies, which he found to majorly support its postulations. However, he still admits that a lot of criticism still holds merit. Schmidt himself amended the hypothesis after the criticism; changing his original position of noticing being essential for language learning to it having a facilitative role (Hulstijn & Schmidt, 1994).

Input enhancement is also closely related to the noticing hypothesis and provides techniques that are used to promote noticing (Smith, 1993). Different techniques are used to bring attention to the target features, including highlighting, underlining and coloring.

3.2.3 Involvement Load Hypothesis

The involvement load hypothesis was proposed by Hulstijn and Laufer (2001) as a method to gauge how effective a technique is in teaching vocabulary. It used cognitive processing as a measure of the effectiveness of a technique. It is essentially a motivational-cognitive model, which considers three factors as major players in learning vocabulary: need, search, and evaluation. It believes that the more these factors will be involved in these activities, the more their effectiveness will increase (Yanagisawa & Webb, 2022).

Yanagisawa and Webb's (2022) meta-analysis of various studies shows that learners learned more items from activities that involved a higher Involvement Load. It further revealed that the presence of these factors might affect learning in different ways; however, there was a marked increase in performance no matter the kind of activity. It does, however, give way to the discussion that different components of Involvement Load may be used to influence in different ways.

3.2.3.1 Need.

According to Hulstijn and Laufer (2001), need is the motivational component of the model and is the source of the drive to learn a specific linguistic feature. Need depends on the source of motivation, which is weak when it is imposed by external means and strong when it is intrinsic. Yanagisawa and Webb (2022) also expand upon

the intensity of the presence of need in terms of the source of its arousal. According to them, it is absent when there is no need for a word in order to complete an activity. It is moderate when the direction for learning the word comes from an external source such as an instructor or a teacher. It finds its strongest presentation when the learner him/herself wants to learn or use the word, leading to them paying the most attention and/or work the hardest.

The motivational nature of need is among the reasons why DDL can be effective as a language teaching method, since learners find interacting with the corpus and finding rules by themselves more interesting.

3.2.3.2 Search.

The second component is search, which is the cognitive part of the hypothesis. It is related to the attempt to find a form in the target language using some source for your need (Hulstijn & Laufer, 2001). Search can be either absent or present. It is completely absent when a linguistic item or word is directly provided by the teacher to the learner, eliminating any need to search for the word or its meaning. Search exists when learners have to actively try to find the meaning of a word, leading them to learn it. An example of such an activity would be a learner reading a text and consulting a dictionary to find out the meanings of words they do not understand (Yanagisawa & Webb, 2022).

3.2.3.3 Evaluation.

Evaluation is also a cognitive component of the Involvement Load hypothesis and is related to the analysis and decision-making process that whether or not the found word is suitable for the context (Hulstijn and Laufer, 2001). According to Yanagisawa and Webb (2022), evaluation is absent when the learner does not need to evaluate the need or use of a word in a certain context. This can be most evident in the presentation step of the PPP method, where the need for evaluation is eliminated because the word is provided by the teacher. Evaluation is moderate when context is available for a word. This can be found in activities like fill-in-the-blanks. Finally, evaluation is the strongest when someone has to make a sentence using a word. This can be strong in the production step of the PPP method.

Hulstijn and Laufer (2001) discuss that the higher these factors are in a vocabulary teaching technique, the more success it will find in being able to teach the

vocabulary. Every activity may have different levels of different components of the Involvement Load and its effectiveness may vary depending on them. Data Driven Learning by its nature provides the most elements where the Involvement Load will be the highest. As we have discussed previously, DDL involves learners interacting with the language data itself either on computers or in the form of printed handouts, where they try to understand the meaning and use of the linguistic item by evaluating the context. In this way, it presents features that may be high on need, search and evaluation. The components of Involvement Load also provide some argument for the noticing hypothesis. Higher levels of need and search may result in increased noticing, which amplifies language learning.

3.2.4 DDL and Constructivism

Constructivism is among the most important theories underpinning the Data Driven Learning approach. According to Flowerdew (2015), it refers to an educational philosophy that considers the acquisition of knowledge, or in this case language, to be a dynamic process. Learners drive their own learning by engaging with language and creating meaning for themselves. This process builds meanings by employing higher order cognitive skills. This inductive process leads to the formation of concepts the minds of the learners.

As Flowerdew (2015) discusses, these postulations of Constructivism have significant similarities to the assumptions underlying DDL. Learners interact with data; in order to understand this data, they apply their problem-solving skills and previous knowledge to contextualize it and break it down and then assimilate it within their knowledge base. These cognitive processes are shared by both the DDL approach and Constructivism.

Other researchers have also added to the theory that corpus use can stimulate cognitive skills in learners. According to O'Sullivan (2007), these skills include “predicting, observing, noticing, thinking, reasoning, analyzing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorizing, hypothesizing and verifying” (p.277). Chang (2012) has also conducted a study on cognitive skills and contrary to O’Sullivan (2007), he contends that the application of higher-order skills, such as inference, was lower in frequency than other cognitive skills. However, he also

maintains that corpus does indeed stimulate cognitive skills, such as making sense and exploring.

3.3 Research Design

The study follows a mixed-method research design. It relies on both quantitative and qualitative methods of data collection. Both kinds of data are obtained simultaneously from the text. It mixes both quantitative and qualitative data using a Convergent Mixed Methods Design. However, it differs from the most popular forms of this design. Qualitative data is collected from the text and then the data is assigned a score and studied quantitatively. This method of analyzing data is called Data Transformation (Creswell & Creswell, 2017). This kind of research design is used to provide a more in-depth look into how Data-Driven Learning affects the acquisition of discourse markers and to identify any patterns beyond the overall ability to use them.

The quantitative part of the study is quasi-experimental. This is due to the fact that the research took place in a real-world, uncontrolled setting where the variables could not be completely controlled. The sampling was completely random; therefore it cannot be called a true experimental study. The study measured the proficiency of learners in using DMs before and after a treatment using a pre-test, a post-test and a delayed post-test. There was a delay of 3 months between the post-test and the delayed post-test, which was employed to gauge the long-term retention of concepts, as it provides insights into the durability of the learning (Schmitt, 2010). The delay period between the two post-tests varies between different studies and there is no standard regarding the delay (Schmitt, 2010), however, the three month delay for the current study was decided on the basis of the quarterly nature of sessions in schools. The empirical data from the quantitative part as well as the quantification of qualitative data using data transformation enabled the researcher to draw out the effects of the treatment procedure of this study.

3.3.1 Pilot Study

In order to study the feasibility of the study and to streamline the lessons, a pilot study was conducted. The subjects of the pilot study were 15 learners of The PEACE Group of Schools & Colleges Abbottabad from grades 11 and 12. The learners had been recently promoted to their grades and had just started the school

year. The learners were administered a pre-test in which they had to write small essays on various subjects. The learners were then divided into two experimental and one control group with 5 learners in each. Out of the two experimental groups, one was administered with soft DDL while the other was treated with hard DDL. Meanwhile, the control group received education through the PPP method, which is traditionally used in classrooms.

3.3.1.1 Results of the Pilot Study.

For the control group, out of the 5 subjects, only 2 used any kind of discourse markers in the small essays they wrote in the pre-test. After teaching the learners using the PPP method, the learners showed slight improvement. Although the learners who previously used at least one discourse marker did not show much improvement, 2 learners, who had not used a discourse marker previously, now used at least one in the post-test. Looking at the standard deviation of the scores, it also decreased from 0.8 to 0.6, which is encouraging. The overall improvement in score was 40%.

Table 1 – Control Group

Test	Learner 1	Learner 2	Learner 3	Learner 4	Learner 5
Pre-test	0	2	1	0	0
Post-test	0	2	1	1	1

Coming to soft DDL, the learners in this group had higher scores in the pre-test than the control group by using a total of 6 discourse markers in their essays. They, however, improved more than the control group by showing 50% improvement, bringing the total number of discourse markers used to 12. The standard deviation in the pre-test was 1.16, which was reduced to 1.01 in the post-test.

Table 2 – Soft Experimental Group

Test	Learner 1	Learner 2	Learner 3	Learner 4	Learner 5
Pre-test	1	0	2	0	3
Post-test	1	2	3	2	4

Finally, coming to hard DDL, the subjects showed 52% improvement in their scores, while reducing the standard deviation from 0.97 to 0.8.

Table 3 – Hard Experimental Group

Test	Learner 1	Learner 2	Learner 3	Learner 4	Learner 5
Pre-test	2	1	2	2	4
Post-test	4	5	4	4	6

3.3.1.2 Discussion of the Pilot Study.

The results of the pilot study are discussed in this section, which indicate that the subjects in all groups improved their scores in the use of discourse markers. This is evidence for the inherent benefit of instruction of DMs without taking the comparison of the two groups into consideration. DDL was found to provide an advantage in learning DMs in comparison with the traditional PPP method, as the subjects in both experimental groups showed over 10% more improvement in their use of discourse markers. The subjects treated with the hard version of DDL showed a little more (2%) improvement than the soft version; however, the difference cannot be considered significant. The difference between the control group and the experimental groups however is more notable and provides reason for further research.

The pilot study also highlighted the need for a change in the testing instrument to a way that directly prompts the learners to use discourse markers instead of only relying on the essays produced by the learners. The answer to this might be found in the methodology used by Boulton (2010) who used questionnaires to provide additional data.

To sum up, the pilot study, albeit small, provided enough indication that both versions of DDL do bring more improvement in the use of discourse markers in comparison with the PPP method. At the same time, it has also provided information to improve the lessons as well as the methodology.

3.3.2 Rationale for the Research Site

The site for research was Modernage Public School & College, Abbottabad. The school ranks among the top schools in the city, attracting mostly learners from the middle class to upper middle economic class. There were two reasons for this school to be chosen. The first reason is that the hard version of DDL relies on computers for the learners to explore the corpus. The research site had a fully

equipped computer lab which was suitable for the purposes of the current study. Secondly, since DDL has mostly been tested in high-level learners, it was believed that learners from a school with better resources will be able to work better with the research treatment.

3.4 Data Collection

The below section details the various elements of data collection for the subject study:

3.4.1 Population

The choice of population for this study was based on the research gap found during the review of relevant literature. An examination of the available research on DDL reveals that there is very little work on the application of DDL at lower levels. Most studies focus on higher level learners, especially learners in universities, because of the perceived complexity of learning through a DDL approach as well as the non-availability of computer equipment for learners at lower levels. This can be observed in Bao's (2021) meta-study on research on DDL, in which we can see that most of these studies evade learners on lower levels, especially in schools. The population for this study comprised students of higher secondary class (12^{th} grade) in Modernage Public School and College Abbottabad. The reasons for choosing learners in higher secondary grades are in line with previous studies, which prefer higher level learners due to the factors mentioned above. The above-mentioned school was chosen because it is a higher-income school and is equipped with sufficient facilities to enable the current study.

3.4.2 Sampling

Purposive sampling was used for the current study. The reason for purposive sampling was that the learners needed to be from schools having good computer equipment to enable the computer-based portion of the study to take place.

3.4.3 Instrument

This research used a pre-test, a post-test and a delayed post-test. The tests comprised two instruments: a sentence-making exercise and an essay. Learners were asked to make 3 sentences for each of the discourse markers in the sentence-making

exercise, which served to directly test their understanding of the different uses of each DM. They were verbally instructed to try to make their sentences as diverse as they could. Moreover, learners had to write an essay at each stage of tests, where they were told to write an essay on a level-appropriate topic provided by the researcher.

The reason for using a sentence-making exercise in the research is that a lot of the times, the learners have the knowledge and understanding to use a discourse marker, however, either the opportunity does not present itself in the context of the text, or they understand the rules but have not yet started to use them actively in their writing. Thus, this method allowed the researcher to directly test the learners' knowledge of discourse markers, by asking them to make sentences using the discourse markers.

Finally, all the three tests–the pre-test, the post-test and the delayed post-test–were evaluated and measured through two methods: quantitative analysis and qualitative analysis turned into quantitative data using data transformation. The quantitative analysis took place by measuring the number of correct usages in the sentence making exercises and counting the number of DMs used in the essays. The qualitative analysis was conducted to determine the variety of usage of DMs in the sentence making exercises. This was accomplished by referring to a key developed from popular dictionaries such as Cambridge, Collins and Merriam Webster and the corpus developed for the current study. The qualitative data, however, was converted into quantitative data by means of Data Transformation (Creswell & Creswell, 2017). These scores were, then, compared across different tests to obtain results.

3.4.4 Materials

A corpus was developed for the purposes of the study. It comprised 51 appropriate grade-level texts. These were obtained from Higher Secondary Level English textbooks from Khyber Pakhtunkhwa Book Board, Peshawar, as well as the Punjab Curriculum and Textbook Board, Lahore. The texts that were used to build the corpus for both of the DDL methods are listed below:

Table 4 – Texts Used in Corpus

S. No.	Text	Author
1.	Button, Button	Richard Matheson
2.	Clearing in the Sky	Jesse Stuart
3.	Dark They were, and Golden Eyed	Ray Bradbury
4.	Thank you, M'am	Langston Hughes
5.	The Piece of String	Guy de Maupassant
6.	The Reward	Lord Dunsany
7.	The Use of Force	William Carlos Williams
8.	The Gulistan of Sa'di	Sheikh Sa'di
9.	The Foolish Quack	Folk Tale
10.	A Mild Attack of Locusts	Doris Lessing
11.	I Have a Dream	Martin Luther King
12.	The Gift of the Magi	O. Henry
13.	God be Praised	Ahmed Nadeem Qasmi
14.	Overcoat	Ghulam Abbas
15.	The Angel and the Author – and Others	Jerome K. Jerome
16.	The Dying Sun	Sir James Jeans
17.	Why Boys Fail in College	Herbert E. Hawkes
18.	End of Term	David Daiches
19.	On Destroying Books	J.C. Squire
20.	The Man who was a Hospital	Jerome K. Jerome
21.	My Financial Career	Stephen Leacock
22.	China's Way to Progress	Galeazzo Santini
23.	Hunger and Population Explosion	Anna McKenzie
24.	The Jewel of the World	Philip K. Hitti
25.	First Year at Harrow	Sir Winston S. Churchill
26.	Hitch-hiking across the Sahara	G. F. Lamb
27.	Sir Alexander Fleming	Patrick Pringle
28.	Louis Pasteur	Margaret Avery
29.	Mustafa Kamal	Wilfrid F. Castle
30.	Quaid e Azam Speech	Muhammad Ali Jinnah

31.	His First Flight	Liam O' Flaherty
32.	From Mother… With Love	Zoa Sherburne
33.	It's Country For Me	Patricia Demuth
34.	Wasteland	Marya Mannes
35.	The White Lamb	Sero Khanzadian
36.	The Importance Of Family	Sam Keen
37.	The Blanket	Floyd Dell
38.	School Vs Education	Russell Baker
39.	Drug Abuse in the Youth of Pakistan	Waheedullah
40.	Quaid's Address To The Constituent Assembly	Muhammad Ali Jinnah
41.	I Have a Dream!	Martin Luther King Jr.
42.	Glory And Hope	Nelson Mandela
43.	A Man Should Never Leave His Post	TuPeng-Cheng
44.	Determination	Anonymous
45.	The Man Who Planted Trees	Jean Giono
46.	The Archaeological Treasures Of Pakistan	Herbert Feldman
47.	Gender Inequality Is Detrimental To The Society	No credited author
48.	Renaissance	Grace Ciaverlla and Angelo Calandra
49.	The Merchant of Venice	Shakespeare
50.	King Lear	Shakespeare
51.	Progress	St. John Green Ervine

3.4.4.1 Building a Corpus.

The corpus was built by taking texts of the stories mentioned in the previous section and pasting them into a text file. The corpus was, then, analysed using a software called AntConc (Version 4.0.0). It is a freeware software providing concordancing ability as well as many other corpus-related features. It was used to analyse the text document serving as the corpus. The total number of words in the corpus was 97,230.

3.4.5 Selection of Discourse Markers

Discourse markers to be taught were selected based on the frequency of their occurrence in the prepared corpus. They can be broadly classified into four categories, including referential, structural, interpersonal, and cognitive, based on the classification driven from Fung and Carter (2007). However, as Maschler and Schiffrin (2015) note, a DM does not necessarily adhere to a single category and may behave differently in various contexts. The DMs for this study have, therefore, not been divided.

The number of DMs chosen for this study was based on the availability of sufficient hits in the corpus as well as previous studies, which have taught a similar number of discourse markers. For example, Jones, (2011) in his doctoral study taught 20 DMs to a group of 36 participants, Kamali and Noori (2015) in their study taught 10 DMs to a group of 60 participants, Alraddadi (2016) taught 21 DMs in his study. The DMs selected for this study are the following along with their frequencies in the corpus:

1. Also – 68 hits
2. Because – 73 hits
3. But – 535 hits
4. However – 41 hits
5. I think – 25 hits
6. Instead – 29 hits
7. Now – 219 hits
8. Oh – 47 hits
9. Perhaps – 29 hits
10. Really – 45 hits
11. So – 304 hits
12. Therefore – 15 hits
13. Then – 240 hits
14. Though – 43 hits
15. Well – 131 hits

3.4.6 Procedure

The research consisted of two experimental groups and one control group. Experimental group No. 1 was taught using the hard version of DDL (based on computers), experimental group No. 2 was taught using the soft version of DDL (paper-based) and the control group was taught using a traditional teaching method, called the Presentation, Practice, Production (PPP) method. There is only one control group for both of the groups. The reason for such an arrangement is that they are both set in the same context. Since the population of the sample is the same, there is no need for two separate control groups. Each group was taught for 8 lessons, with each being 40 minutes. The first lesson was for the purpose of orientation, in which the learners were introduced to the study as well as DMs. Lesson plans were drawn for each of the 7 classes left. The length of treatment delivered in this study is comparable to similar studies. For example, Rahimi and Riasati (2012) taught 40 participants for 5 lessons, each spanning 20 minutes; Kamali and Noori (2015) delivered 10 lessons, 30 minutes each, however, notably the number of DMs in their study was more than the current study, making the time spent per DM greater in this study; Hernández (2011) gave two 50 minute lessons to 91 participants and so on.

The learners in the hard DDL or Hard Experimental Group were allowed to explore the corpus freely. The corpus was edited beforehand to remove any undesirable use of language that may appear in the concordance for the participants. For the soft DDL, since the space on the handouts was limited, the researcher picked out instances of the use which were the most diverse as well as most accurately provided an example of the use of the DM.

The research originally aimed to collect at least 120 samples; 40 for each group. Consent was taken from a total of 160 learners, however, only 110 learners returned both the completed sentence-making exercises and the essays for all three stages, with 33 (Control Group), 31 (Hard DDL Group) and 46 (Soft DDL Group) learners returning both of these. To equalize the number of samples across all the groups, 30 learners for every group were chosen for each group for analysis and extra samples were chosen to be left out at random.

3.4.7 Lesson Plans

Lessons were chosen based on the selection of DMs as explained above. The choice of the number of DMs to be taught in each lesson was made on the basis of their perceived difficulty. There were three sets of lesson plans. The first set of lesson plans was for the control group, the second was for the Hard Experimental Group (HEG) and the third was for the Soft Experimental Group (SEG). The details of these lesson plans are given below:

3.4.7.1 Lesson Plans for the Control Group.

3.3.7.1.1 Orientation.

Objectives:

1. To familiarize the learners with the research topic
2. To familiarize the learners with the research objectives
3. To describe the variables of the research
4. To explain ethical considerations and confidentiality
5. To familiarize the learners with the research software

Duration of the Lesson: 40 minutes

The first lesson was the orientation of the learners regarding the research and its purposes. The researcher introduced himself and his parent institute for research. The learners were then oriented about the purpose of the research. All the variables of the research were defined and explained with examples. The researcher allowed for questions to be asked. At the end, consent was taken from the participants. Ethical considerations and confidentiality terms were explained.

3.4.7.1.2 Lesson Plan 1: However & Therefore.

Objectives

1. To teach the learners the use of the DM “However” using the PPP method
2. To teach the learners the use of the DM “Therefore” using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

However

1. Used to show contrast (*Collins English Dictionary*, 2024)
    a. E.g. “This was not an easy decision. It is, however, a decision that we feel is dictated by our duty.” (*Collins English Dictionary*, 2024)
    b. E.g. “Some of the food crops failed. However, the cotton did quite well.” (*Collins English Dictionary*, 2024)
2. In the sense of despite this (*However*, 2024)
    a. E.g. “This is one possible solution to the problem. However, there are others.” (*However*, 2024)
    b. E.g. “There may, however, be other reasons that we don't know about.” (*However*, 2024)
3. In the sense of ‘On the other hand’ (*Definition of HOWEVER*, n.d.)
    a. E.g. “however, I think I'd better not” (*Definition of HOWEVER*, n.d.)

Therefore

1. To introduce a logical result (*THEREFORE | Collins English Dictionary*, 2024)
    a. E.g. “I think, therefore I am” (*THEREFORE | Collins English Dictionary*, 2024)
    b. E.g. “they heard the warning on the radio and therefore took another route” (*THEREFORE | Collins English Dictionary*, 2024)
2. To provide a conclusion (*THEREFORE | Collins English Dictionary*, 2024)
    a. E.g. “Muscle cells need lots of fuel and therefore burn lots of calories.” (*THEREFORE | Collins English Dictionary*, 2024)
    b. E.g. “those people have their umbrellas up: therefore, it must be raining.” (*THEREFORE | Collins English Dictionary*, 2024)
3. In the sense of ‘for that reason’ (*Therefore*, 2024)
    a. E.g. “We were unable to get funding and therefore had to abandon the project.” (*Therefore*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.3 Lesson Plan 2: Also, now & then.

Objectives

1. To teach the learners the use of the DM "Also" using the PPP method
2. To teach the learners the use of the DM "Now" using the PPP method
3. To teach the learners the use of the DM "Then" using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

Also

1. To provide more information (*ALSO Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. "She has a reputation for brilliance. Also, she is a good communicator." (*ALSO Definition and Meaning | Collins English Dictionary*, 2024)
2. To show similarity (*ALSO Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. "His father, also a top-ranking officer, had perished during the war." (*ALSO Definition and Meaning | Collins English Dictionary*, 2024)

   b. E.g. “Not only cancer, but also heart and lung disease are influenced by smoking.” (*ALSO Definition and Meaning | Collins English Dictionary*, 2024)
3. In the sense of “in addition” (*Also*, 2024)
   a. E.g. “She's a photographer and also writes books.” (*Also*, 2024)

Now

1. To refer to the present (*NOW Definition and Meaning | Collins English Dictionary*, 2024)
   a. E.g. “She's a widow now” (*NOW Definition and Meaning | Collins English Dictionary*, 2024)
   b. E.g. “She should know that by now.” (*NOW Definition and Meaning | Collins English Dictionary*, 2024)
2. To indicate the time immediately before the present (*Definition of NOW*, 2024)
   a. E.g. “thought of them just now” (*Definition of NOW*, 2024)
3. To show command where the sense of present time is weakened (*Definition of NOW*, 2024)
   a. E.g. “now hear this”(*Definition of NOW*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.4 Lesson Plan 3: Really, Perhaps.

Objectives

1. To teach the learners the use of the DM “Really” using the PPP method
2. To teach the learners the use of the DM “Perhaps” using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

Really

1. To put emphasis (*REALLY Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “I really do feel that some people are being unfair.” (*REALLY Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “You know, we really ought to get another car.” (*REALLY Definition and Meaning | Collins English Dictionary*, 2024)
2. To express surprise (*Really*, 2024)
    a. E.g. “"I'm getting married to Fred." "Really? When?"” (*Really*, 2024)
3. To express disbelief or doubt (*REALLY Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Do you really think he would be that stupid?” (*REALLY Definition and Meaning | Collins English Dictionary*, 2024)

Perhaps

1. To show uncertainty (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “In the end they lose millions, perhaps billions.” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “It was bulky, perhaps three feet long and almost as high.” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
2. To make opinions polite (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Perhaps the most important lesson to be learned is that you simply cannot please everyone.” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)

   b. E.g. “His very last paintings are perhaps the most puzzling.” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)

3. To make requests (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
   a. E.g. “Perhaps I may be permitted a few suggestions.” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)
   b. E.g. “Perhaps if you rang me when you got back to your office?” (*PERHAPS Definition and Meaning | Collins English Dictionary*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.5 Lesson Plan 4: Though & Instead.

Objectives

1. To teach the learners the use of the DM “Though” using the PPP method
2. To teach the learners the use of the DM “Instead” using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

Though

1. To introduce a contrasting clause (*THOUGH Definition and Meaning | Collins English Dictionary*, 2024)

   a. E.g. “The film was exactly how I had pictured it, though I think Gale should have had a bigger part.” (*THOUGH Definition and Meaning | Collins English Dictionary*, 2024)
   b. E.g. “The rest of his 'team' are simply assistants, though all very good at what they do.”
2. In the sense of “despite the fact” (*Though*, 2024)
   a. E.g. “She hasn't called, even though she said she would.” (*Though*, 2024)
3. In the sense of “in spite of the possibility that” (*Definition of THOUGH*, n.d.)
   a. E.g. “though I may fail, I will try” (*Definition of THOUGH*, n.d.)

Instead

1. To provide an alternative (*Instead*, 2024)
   a. E.g. “We went by train instead of by car.” (*Instead*, 2024)
2. To replace something (*Instead*, 2024)
   a. E.g. “There's no coffee - would you like a cup of tea instead?” (*Instead*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.6 Lesson Plan 5: So & But.

Objectives

1. To teach the learners the use of the DM “So” using the PPP method
2. To teach the learners the use of the DM “But” using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

So

1. To indicate result (*Also*, 2024)
    a. E.g. “the acoustics are good, so every note is clear” (*Also*, 2024)
2. In the sense of “in order to” (*Also*, 2024)
    a. E.g. “be quiet so he can sleep” (*Also*, 2024)
3. To introduce a sentence (*Also*, 2024)
    a. E.g. “so here we are” (*Also*, 2024)

But

1. To show contrast (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “He not only wants to be taken seriously as a musician, but as a poet too.” (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
2. To change the subject (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “They need to recruit more people into the prison service. But another point I'd like to make is that many prisons were built in the nineteenth century.” (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
3. In the sense of ‘except’ (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Europe will be represented in all but two of the seven races.” (*BUT Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “He didn't speak anything but Greek.” (*BUT Definition and Meaning | Collins English Dictionary*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.7 Lesson Plan 6: Well & Because.

Objectives

1. To teach the learners the use of the DM "Well" using the PPP method
2. To teach the learners the use of the DM "Because" using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

Well

1. To indicate the start of an utterance (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. "Well, you go get yourselves some breakfast." (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. "Well, I don't like the look of that." (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
2. To make a statement more polite (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. "Well, let's wait and see." (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. "Well, I thought she was a bit unfair about me." (*WELL Definition and Meaning | Collins English Dictionary*, 2024)

3. To indicate the end of a conversation (*WELL Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Well, thank you for speaking with us.” (*WELL Definition and Meaning | Collins English Dictionary*, 2024)

Because

1. To provide reason (*BECAUSE Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “He is called Mitch, because his name is Mitchell.” (*BECAUSE Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “Because it is an area of outstanding natural beauty, you can't build on it.” (*BECAUSE Definition and Meaning | Collins English Dictionary*, 2024)
2. In the sense “as a result of” (*Because*, 2024)
    a. E.g. “The trip was canceled because of bad weather.” (*Because*, 2024)
3. To provide explanation of a statement (*BECAUSE Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Maybe they didn't want to ask questions, because they rented us a room without even asking to see our papers.” (*BECAUSE Definition and Meaning | Collins English Dictionary*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.1.8 Lesson Plan 7: I think & Oh.

Objectives

1. To teach the learners the use of the DM “I think” using the PPP method
2. To teach the learners the use of the DM “Oh” using the PPP method

Material: Whiteboard and notebooks

Duration of the Lesson: 40 minutes

Presentation

The first part of the lesson involved the presentation of the concepts. The use and meanings as the examples of each of the DMs were taken from popular dictionaries:

I think

1. To indicate thinking process (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “I think I'll go home and have a shower.” (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “Time for a cup of coffee, I think.” (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)
2. To not be rude or forceful (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Thanks, but I think I can handle it.” (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “This is, I think, much, much more important.” (*I THINK Definition and Meaning | Collins English Dictionary*, 2024)

Oh

1. To indicate surprise/excitement (*OH Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “Oh, I'm so glad you're here.” (*OH Definition and Meaning | Collins English Dictionary*, 2024)
    b. E.g. “’Oh!' Kenny blinked. 'Has everyone gone?'” (*OH Definition and Meaning | Collins English Dictionary*, 2024)
2. To provide a response (*OH Definition and Meaning | Collins English Dictionary*, 2024)
    a. E.g. “'You don't understand!'—'Oh, I think I do, Grace.'” (*OH Definition and Meaning | Collins English Dictionary*, 2024)

   b. E.g. “Would you like me to phone and explain the situation?'—'Oh, would you?'” (*OH Definition and Meaning | Collins English Dictionary*, 2024)
3. To indicate disappointment/dismay (*Oh*, 2024)
   a. E.g. “Oh dear, what a mess!” (*Oh*, 2024)
   b. E.g. “Oh no, I left my umbrella behind!” (*Oh*, 2024)

Practice

The learners were asked to make sentences with DMs with the assistance of the researcher.

Production

The learners were tasked to make 5 sentences for each of the DMs in their notebooks without any help.

3.4.7.2 Lesson Plans for HEG:

3.4.7.2.1 Orientation.

Objectives

6. To familiarize the learners with the research topic
7. To familiarize the learners with the research objectives
8. To describe the variables of the research
9. To explain ethical considerations and confidentiality
10. To familiarize the learners with the research software

Duration of the Lesson: 40 minutes

The first lesson was the orientation of the learners regarding the research and its purposes. The researcher introduced himself and his parent institute for research. The learners were then oriented about the purpose of the research. All the variables of the research were defined and explained with examples. The researcher allowed for questions to be asked. In the end, consent was taken from the participants. Ethical considerations and confidentiality terms were explained.

The participants were also provided a demonstration of the software that will be used for the research and how to use it.

3.4.7.2.2 Lesson Plan 1: However & Therefore.

Objectives

3. To help learners discover the use of the DM “However” independently
4. To help learners discover the use of the DM “Therefore” independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

The first lesson was focused on two DMs in particular: However and therefore. The learners were guided to discover the meanings of the DMs on their own. This was done by them by seeing how the DMs affected the statements immediately behind and in front of them. The researcher (in this case the instructor too) asked questions about how the DMs were used and pointed the learners towards making comparisons of different uses. These questions included questions as such if they could see the same meaning from the use of the DM in another place that they found in one instance. If they could not apply the same meaning, they were encouraged to identify how they were different. The learners were further encouraged to click on the DMs and look at how they found their place in the broader context of the paragraph.

The learners were asked to write down the examples where they thought the DMs were used in a different way than the others. In the end, they were asked to share their own explanations that they made and compare them with each other. Both DMs were allotted approximately 20 minutes each, with the last 5 minutes of these allocated for the comparison process.

3.4.7.2.3 Lesson Plan 2: Also, now & then.

Objective

1. To help learners discover the use of the DM “Also” independently
2. To help learners discover the use of the DM “Now” independently
3. To help learners discover the use of the DM “Then” independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

The second lesson was focused on three DMs due to their relative ease. These included ‘also’, ‘now’ and ‘then’. Each DM was allotted 10-15 minutes. Similar to the first lesson, the researcher helped the learners in discovering the meanings of the DMs independently by seeing what kind of relationship they hold with their context. They were able to explore further context by clicking on the keyword and seeing the broader context of the use of the word. The instructor guided the learners by asking questions about the use of the DMs, which allowed them to compare how the DMs were used in different instances and if they could apply the same meaning in different places.

During each slot, the learners also compared how their interpretations of the meanings differed, allowing their learning to be more enriched.

3.4.7.2.4 Lesson Plan 3: Really, Perhaps.

Objectives

1. To help learners discover the use of the DM “Really” independently
2. To help learners discover the use of the DM “Perhaps” independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

For the third lesson, the researcher focused on two DMs: Really and Perhaps. The allotted time for each DM was 20 minutes with the last 5 minutes of each reserved for comparison of the findings of the learners. The instructor facilitated the learners in their exploration of the use of the DMs in their respective contexts. The AntConc software allowed the learners to delve deeper into the context of the DM by clicking on the keywords to see the wider context. The instructor again asked questions to guide the learners through the comparison of the DMs across various instances and in assessing how they can be applied. Finally, at the end, the learners repeated the comparison exercise to help each other flesh out their findings.

3.4.7.2.5 Lesson Plan 4: Though & Instead.

Objectives

1. To help learners discover the use of the DM “Though” independently
2. To help learners discover the use of the DM “Instead” independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

The fourth lesson targeted the DMs: Though and instead. The allocation of time was maintained with 20 minutes for each DM with the last part of them being time for comparison among the participants of the group. The process of finding the meanings of DMs within their context while also exploring the broader context was again followed in this lesson. This was further facilitated by the researcher's questions which directed the attention of the learners towards the peculiarities of use of the DMs. Notes were shared at the end of each exercise.

3.4.7.2.6 Lesson Plan 5: So & But.

Objectives

1. To help learners discover the use of the DM "Though" independently
2. To help learners discover the use of the DM "Instead" independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

The fifth lesson explored the meaning of the DMs 'so' and 'but'. While the DMs looked relatively easy, the questions of the researcher directed the learners towards the nuances of the use of the DMs. The comparison between different examples of the DMs found with the help of the concordancer helped further pronounce the various uses. Similar to previous attempts, learners were also facilitated to find out more about how the DMs are treated in the context of paragraphs. The learners continued the practice of sharing their findings to enrich their learning. The allocated time remained to be 20 minutes of each DM with the last 5 minutes of each being specified for comparison.

3.4.7.2.7 Lesson Plan 6: Well & Because.

Objectives

1. To help learners discover the use of the DM "Well" independently
2. To help learners discover the use of the DM "Because" independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

The sixth lesson delved into the use of the DMs: well and because. The instructors asked the learners to explore the usage of the DMs using the AntConc software and urged them to explore the usage of the above DMs. The instructor, then, asked questions to help the learners along their research into the DMs while they explored the instances within the concordance as well as their broader context. At the end of both of the 20-minute slots, the learners shared their findings with each other to help them verbalize and flesh out their findings.

3.4.7.2.8 Lesson Plan 7: I think & Oh.

Objectives

1. To help learners discover the use of the DM "I think" independently
2. To help learners discover the use of the DM "Oh" independently

Material: Computer, AntConc Software and the study corpus

Duration of the Lesson: 40 minutes

During the last lesson for the group, the learner sought to facilitate the learners to research the uses and meanings of the DMs 'I think' and 'Oh'. This was again done in three key ways: (1) by looking at the different instances of use of the DMs, (2) by looking at the broader context in the paragraphs and (3) by asking questions to direct attention towards certain features and peculiarities. The learners shared their findings with each other at the end of the designated 20-minute slots to further deepen their knowledge.

3.4.7.3 Lesson Plans for SEG:

3.4.7.3.1 Orientation.

Objectives

1. To familiarize the learners with the research topic
2. To familiarize the learners with the research objectives
3. To describe the variables of the research
4. To explain ethical considerations and confidentiality
5. To familiarize the learners with the research handouts

Duration of the Lesson: 40 minutes

The first lesson was the orientation of the learners regarding the research and its purposes. The researcher introduced himself and his parent institute for research. The learners were then oriented about the purpose of the research. All the variables of the research were defined and explained with examples. The researcher allowed for questions to be asked. At the end, consent was taken from the participants. Ethical considerations and confidentiality terms were explained.

The participants were provided a sample of the handouts that they would receive during each lesson and were given a demonstration of how the process will go.

3.4.7.3.2 Lesson Plan 1: However & Therefore.

Objectives

1. To help learners discover the use of the DM “However” independently
2. To help learners discover the use of the DM “Therefore” independently

Material: Research handouts

Duration of the Lesson: 40 minutes

The first lesson was focused on two DMs in particular: However and therefore. The learners were guided to discover the meanings of the DMs on their own. This was done by them by seeing how the DMs affected the statements immediately behind and in front of them. The researcher asked questions about how the DMs were used and pointed the learners towards making comparisons of different uses. These questions included questions as such if they could see the same meaning from the use of the DM in another place that they found in one instance. If they could not apply the same meaning, they were encouraged to identify how they were different.

The learners were asked to write down the examples where they thought the DMs were used in a different way than the others. In the end, they were asked to share their own explanations that they made and compare with each other. Both DMs were allotted approximately 20 minutes each, with the last 5 minutes of these allocated for the comparison process.

3.4.7.3.3 Lesson Plan 2: Also, now & then.

Objectives

1. To help learners discover the use of the DM "Also" independently
2. To help learners discover the use of the DM "Now" independently
3. To help learners discover the use of the DM "Then" independently

Material: Research handouts

Duration of the Lesson: 40 minutes

The second lesson was focused on three DMs due to their relative ease. These included 'also', 'now' and 'then'. Each DM was allotted 10-15 minutes. Similar to the first lesson, the researcher helped the learners in discovering the meanings of the DMs independently by seeing what kind of relationship they hold with their context. The instructor guided the learners by asking questions about the use of the DMs, which allowed them to compare how the DMs were used in different instances and if they could apply the same meaning in different places.

During each slot, the learners also compared how their interpretations of the meanings differed, allowing their learning to be more enriched.

3.4.7.3.4 Lesson Plan 3: Really, Perhaps.

Objectives

3. To help learners discover the use of the DM "Really" independently
4. To help learners discover the use of the DM "Perhaps" independently

Material: Research handouts

Duration of the Lesson: 40 minutes

For the third lesson, the researcher focused on two DMs: Really and Perhaps. The allotted time for each DM was 20 minutes with the last 5 minutes of each reserved for comparison of the findings of the learners. The instructor facilitated the learners in their exploration of the use of the DMs in their respective contexts by comparing their uses in the examples listed in the handouts. The instructor again asked questions to guide the learners through the comparison of the DMs across various instances and in assessing how they can be applied. Finally, at the end, the learners repeated the comparison exercise to help each other flesh out their findings.

3.4.7.3.5 Lesson Plan 4: Though & Instead.

Objectives

3. To help learners discover the use of the DM “Though” independently
4. To help learners discover the use of the DM “Instead” independently

Material: Research handouts

Duration of the Lesson: 40 minutes

The fourth lesson targeted the DMs: Though and instead. The allocation of time was maintained with 20 minutes for each DM with the last part of them being time for comparison among the participants of the group. The process of finding the meanings of DMs within their context was again followed in this lesson. This was further facilitated by the researcher’s questions which directed the attention of the learners towards the peculiarities of use of the DMs. Notes were shared among the participants at the end of each exercise.

3.4.7.3.6 Lesson Plan 5: So & But.

Objectives

3. To help learners discover the use of the DM “Though” independently
4. To help learners discover the use of the DM “Instead” independently

Material: Research handouts

Duration of the Lesson: 40 minutes

The fifth lesson explored the meaning of the DMs ‘so’ and ‘but’. While the DMs looked relatively easy, the questions of the researcher directed the learners towards the nuances of the use of the DMs. The comparison between different examples of the DMs found in the examples on the handouts allowed the learners to discern more nuisances in their use. Similar to previous attempts, learners were also facilitated to find out more about how the DMs are treated in the context of paragraphs. The learners continued the practice of sharing their findings to enrich their learning. The allocated time remained to be 20 minutes of each DM with the last 5 minutes of each being specified for comparison.

3.4.7.3.7 Lesson Plan 6: Well & Because.

Objectives

3. To help learners discover the use of the DM “Well” independently
4. To help learners discover the use of the DM “Because” independently

Material: Research handouts

Duration of the Lesson: 40 minutes

The sixth lesson delved into the use of the DMs: well and because. The instructors asked the learners to explore the usage of the DMs using the provided handouts of the corpus and urged them to explore the usage of the above DMs. The instructor, then, asked questions to help the learners along their research into the DMs. At the end of both of the 20-minute slots, the learners shared their findings with each other to help them verbalize and flesh out their findings.

3.4.7.3.8 Lesson Plan 7: I think & Oh.

Objectives

3. To help learners discover the use of the DM “I think” independently
4. To help learners discover the use of the DM “Oh” independently

Material: Research handouts

Duration of the Lesson: 40 minutes

During the last lesson for the group, the learner sought to facilitate the learners to research the uses and meanings of the DMs ‘I think’ and ‘Oh’. This was done in two key ways: by looking at the different instances of use of the DMs, and by asking questions to direct attention towards certain features and peculiarities. The learners shared their findings with each other at the end of the designated 20-minute slots to further deepen their knowledge.

3.5 Method for Data Analysis

Data was collected by evaluating the samples obtained in the form of sentence making exercises and essays.

3.5.1 Analysis of Sentence Making Exercises

The sentence making exercises were evaluated in two ways: one quantitative and one qualitative. The quantitative part involved marking the uses of DMs out of a score of 3 for each DM (or 45 for all 15 DMs). This was based on the instructions to the learners for making three sentences for each of the DMs. This provided an overall general score for each learner. For the qualitative analysis, a key was developed from the corpus as well as popular dictionaries (Annexure – B). This key was referenced to determine how many diverse ways the DMs were used to make sentences. The DMs were, then, assigned scores based on this determination in a process called Data Transformation (Creswell & Creswell, 2017), which allows quantitative analysis of qualitative data.

The data obtained from both of these processes was analysed using a statistical software called SPSS. The SPSS software provided the ability to run ANOVA and t-tests on the data, which were employed according to the requirements of the research questions.

3.5.2 Analysis of Essays

Essays were analysed to provide quantitative data regarding the use of DMs by the learners in their natural writing. DMs were identified based on the characteristics and types of DMs discussed in the literature review of this study (e.g. Fung & Carter, 2007; Maschler & Schiffrin, 2015). There were two kinds of scores assigned to each essay. A General DMs Score which indicated the total number of DMs used by the participant in his/her essay and a Study Specific DMs Score, which indicated the number of DM used by the learners that were taught as a part of this study.

Similar to the analysis of the sentence making exercises, SPSS was used to run ANOVA and t-test based on the research questions, which provided statistical analysis of the data to be further interpreted.

The results of both of these sources of data were, then, interpreted based on the theoretical framework, drawing inferences about how its different elements may have been operational in explaining the results found in different components of the instruments.

3.6 Validity and Reliability

Validity and reliability are important measures of how useful and trustworthy a research study can be. Maintaining both of them is essential for the results and interpretations of a research study to be a true reflection of reality and be credible enough to add to the scientific canon. Roberts and Priest (2006) believe that the validity of a study lies in making sure that the measurements made in a study align with its objectives. On the other hand, reliability has to do with the generalizability and stability of the study, i.e. if it would be able to produce similar results with similar parameters, if performed in different circumstances. A study may be at risk of exposure to several internal and external threats to its validity, such as the population, participants, settings, research instruments, etc., involved in the study. Adequate measures must be taken in every research study to ensure that both validity and reliability are maintained.

The current study relied on various factors to ensure its validity. The essay writing research instrument is popular for measuring the use of DMs and has been used by various studies (e.g. Aidinlou & Mehr, 2012; Kapranov, 2018; Sadeghi & Kargar, 2014). The topics for the essays were spontaneous and the process of writing the essays was conducted under supervision to ensure they did not copy from online sources. The DMs were identified based on the characteristics and types established by Fung and Carter (2007 and Maschler and Schiffrin (2015). Coming to the sentence making exercise, it was based on the pilot study conducted for this study, which presented the need for a way to directly test for the learners' knowledge of the DMs. These tests were scored based on a key developed using various dictionaries, as found in Annexure B.

The use of two research instruments enabled triangulation of the data, which is a technique through which two data sources are combined to provide more robust, comprehensive and consistent results (Roberts & Priest, 2006). In the current study, both the sentence making exercise and the essay provided data, both

quantitative and qualitative, which supported each other, hence, adding to its trustworthiness. Finally, the results of the study were supported by the results of other studies relevant to the variables of this study, which provided further credibility to them.

3.7 Ethical Considerations

Research is a lengthy process involving interactions with the environment, people and a large amount of decision making about these interactions. Ethical considerations are the moral compass that guides these interactions and decision making. It also dictates how any information is treated and makes sure that neither the environments nor the settings of the study nor the participants are harmed in any way. As Creswell and Creswell (2017) point out they ensure that the rights, needs and values of the participants are protected and allow the researcher to maintain transparency, confidentiality and fairness in the research process. They become even more important when a study is collecting data directly from participants, similar to the subject study.

The ethical considerations of the subject study involved articulating the objectives and procedure of the study to the potential participants, obtaining informed consent and maintaining confidentiality. At the start of the study, the participants were provided a leaflet including the details of the study as well as their rights to privacy. They were also provided verbal clarification and their questions were answered to their satisfaction. Finally, informed consent was obtained from the participants who were willing to join the study. The data obtained from the participants was kept in strict confidentiality by the researcher and was not used for anything other than the purposes of the current study. The data has been anonymised for this study and does not identify or share any personal information of the participants.

3.8 Summary

The above chapter provided the methodology used in the current study, including the population for the study and the samples collected. It provided the theoretical framework on which the study relied. It also discussed the pilot study conducted to refine the current research. The materials and instruments for the purposes of this study were also described in detail. The chapter also laid out the

process of building a corpus and how it was used to inform the methodology. Finally, it detailed the lesson plans that were followed for its purposes.

CHAPTER 4

DATA ANALYSIS AND RESULTS

4.1 Overview

This chapter analyzes the data collected in the current study and presents results in relation with the postulated hypotheses. The presented results are descriptive in nature and are accompanied by a brief discussion in light of the theoretical framework.

4.2 Data

The data for the current study comes from the pre-test, post-test and delayed post-test administered to the participants of the study. Each test consisted of two instruments: the essay and the sentence-making exercise. The samples for the pre-test were collected under the supervision of the researcher. However, the latter two samples were collected by the respective teachers of the learners due to time constraints.

4.2.1 Essays

For the essays, the data was collected by simply counting the number of DMs used in each test. The learners were instructed to write an essay comprising at least 500 words. The topics of the essays were spontaneous and argumentative. The following essay topics were given for each of the tests:

Pre-test – “Impact of Social Media”

Post-test – “Pros and Cons of the Internet”

Delayed Post-test – “Importance of education”

Data was collected by counting the number of DMs used in each of the tests. Another layer of data was collected by specifically looking at the number of DMs in the essays that were taught during the treatment.

4.2.2 Sentence Writing Exercise

A sentence writing exercise was used to assess the conceptual understanding of every specific DM taught during the study. The purpose of this exercise was to

triangulate the data and thus, increasing the validity. It served to directly gauge the level of understanding of the DMs, since sometimes learners know about the meaning and use of a DM, but, they are not able to use them in their language.

Each test was assigned two scores. The first score reflected the times a learner correctly used a DM. The second score was obtained by a process called Data Transformation (Creswell & Creswell, 2017), which converts qualitative data into quantitative data by assigning scores. The second kind of analysis is qualitative because it analyses the different kinds of ways a DM is used. To elaborate, it sees if a DM has been used in more than one way and what kind of relationship it signifies. Since the determination of the different kinds of meanings can be a subjective affair despite having guidelines about the specific meanings that exist, data is considered qualitative. However, since the current study is mainly a quantitative one, the data is transformed into quantitative data by transforming it using this score.

4.3 Statistical Analysis

This section provides the statistical analysis of the obtained data and then tests the hypotheses against the data. The data was analyzed through the software SPSS using an ANOVA test. An ANOVA test was used because the current study includes more than two groups.

4.3.1 Homogeneity of the groups in the pre-test

For the validity of the research, it is essential that the control and experimental groups are similar before the treatment (Creswell, 2012). This can be assessed by comparing the pre-test scores of the participants. Below are the pre-test scores of the participants of each of the groups.

Control Group – Pre-test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Cumulative Score
1.	41	16	6	4	67
2.	28	18	4	2	52
3.	40	25	20	10	95

4.	26	15	5	5	51
5.	42	21	4	3	70
6.	44	20	8	5	77
7.	36	19	14	8	77
8.	15	15	1	1	32
9.	33	20	1	1	55
10.	40	18	11	3	72
11.	31	21	1	1	54
12.	42	19	7	2	70
13.	42	21	6	4	73
14.	37	17	9	7	70
15.	37	16	0	0	53
16.	42	23	7	4	76
17.	39	18	10	8	75
18.	36	19	8	3	66
19.	39	17	9	7	72
20.	40	22	9	5	76
21.	39	19	5	3	66
22.	42	23	14	9	88
23.	25	16	4	4	49
24.	43	22	6	4	75
25.	33	17	7	6	63
26.	19	19	5	5	48
27.	29	20	5	3	57
28.	36	18	8	5	67
29.	37	17	14	10	78
30.	31	13	6	4	54

Mean = 65.93, SD = 13.26 (Cumulative Score)

Soft DDL Experimental Group – Pre-test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Sum of the Score
1.	36	14	15	12	77
2.	41	19	5	5	70
3.	42	20	8	4	74
4.	39	20	3	2	64
5.	39	21	11	6	77
6.	14	14	1	1	30
7.	38	21	12	7	78
8.	39	18	5	3	65
9.	37	18	19	3	77
10.	37	17	5	1	60
11.	38	20	6	2	66
12.	40	21	14	5	80
13.	37	18	11	5	71
14.	41	21	4	4	70
15.	17	13	3	2	35
16.	36	17	5	1	59
17.	27	17	6	4	54
18.	29	18	12	5	64
19.	36	20	7	5	68
20.	36	19	5	5	65
21.	15	12	1	1	29
22.	42	18	13	8	81
23.	41	17	6	4	68
24.	38	18	20	9	85
25.	38	24	8	2	72
26.	40	19	5	5	69
27.	34	20	9	3	66
28.	37	20	7	2	66

29.	34	17	1	1	53
30.	36	18	6	6	66

Mean = 65.3, SD = 13.74 (Cumulative Score)

Hard DDL Experimental Group – Pre-test

Participant No	**Essays – General DMs Score**	**Essays – Study DMs Score**	**Sentences General Score**	**Sentences Meanings Score**	**Cumulative Score**
1.	39	17	3	1	60
2.	35	17	15	10	77
3.	17	15	7	6	45
4.	28	15	13	7	63
5.	39	19	6	3	67
6.	38	19	10	7	74
7.	42	16	4	1	63
8.	16	16	11	7	50
9.	35	19	6	5	65
10.	45	27	7	4	83
11.	39	19	6	6	70
12.	37	22	6	3	68
13.	39	17	5	4	65
14.	40	19	5	1	65
15.	41	21	11	9	82
16.	36	16	4	3	59
17.	37	21	7	4	69
18.	36	19	12	7	74
19.	37	17	1	0	55
20.	17	16	1	1	35
21.	44	22	1	1	68
22.	39	21	5	1	66
23.	42	20	8	2	72
24.	41	28	12	7	88
25.	38	18	4	4	64

26.	44	21	1	1	67
27.	29	14	11	6	60
28.	38	17	9	4	68
29.	28	11	9	6	54
30.	39	19	9	6	73

Mean = 65.63, SD = 10.93 (Cumulative Score)

The following statistics are revealed using an ANOVA test:

Comparison of the Groups – Pre-Test

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	.63	3.28	.98
	Hard	30	.30	3.28	.99
Soft	PPP	30	.63	3.28	.98
	Hard	30	.33	3.28	.99
Hard	PPP	30	.30	3.28	.99
	Soft	30	.33	3.28	.99

The high significance or the P value (0.98 and 0.99) reveals that the difference between the three groups is not significant. The higher the P value, the less significant the difference in the results is. In this case, the P value is significantly higher than the threshold value of 0.05. The low F ratio is also an indicator of the low difference between the three groups in the pre-test, therefore it can be said that the groups were fairly similar before the treatment. The groups are fairly homogenous and the differences are not significant. This provides us a good starting point for this study, because equal statistics in the pre-test will give us a reliable platform for conducting our study.

4.3.2 Null Hypotheses Testing

This section tests all of the null hypotheses one by one in light of the study data. The data allows us to either accept or reject the null hypotheses.

4.3.2.1 Null Hypothesis 1: There is No Significant Difference in the Post-Test Scores of the Three Groups.

The following three tables include the scores of the participants of all three research groups in the post-test of the study.

Control Group – Post-test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Cumulative Score
1.	37	22	12	9	80
2.	42	20	14	8	84
3.	37	14	2	2	55
4.	33	16	4	4	57
5.	43	28	2	3	76
6.	33	15	10	8	66
7.	45	22	5	4	76
8.	42	20	6	6	74
9.	39	19	16	10	84
10.	37	23	12	8	80
11.	30	19	6	5	60
12.	40	22	16	3	81
13.	35	21	5	3	64
14.	39	21	0	0	60
15.	37	22	10	6	75
16.	40	24	9	7	80
17.	21	17	10	7	55
18.	40	22	16	10	88
19.	40	17	8	8	73
20.	40	25	23	9	97
21.	39	22	17	12	90
22.	44	17	8	7	76
23.	38	20	9	5	72

24.	31	17	4	4	56
25.	41	21	12	6	80
26.	39	20	12	6	77
27.	38	21	9	4	72
28.	40	22	6	4	72
29.	31	17	6	5	59
30.	35	16	4	4	59

Mean = 72.60, SD = 11.28 (Cumulative Score)

Soft DDL Experimental Group – Post-test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Sum of the Score
1.	41	21	6	6	74
2.	42	22	18	12	94
3.	39	20	14	7	80
4.	44	25	16	12	97
5.	40	24	21	11	96
6.	44	25	7	5	81
7.	40	21	19	13	93
8.	42	22	9	4	77
9.	37	22	9	5	73
10.	42	2	11	6	61
11.	41	25	8	8	82
12.	40	17	14	9	80
13.	40	20	21	12	93
14.	37	18	11	8	74
15.	44	26	8	7	85
16.	40	21	10	9	80
17.	42	22	11	8	83
18.	40	24	8	5	77
19.	41	21	20	11	93
20.	42	21	12	10	85

21.	45	24	11	10	90
22.	41	19	15	8	83
23.	37	22	12	8	79
24.	40	21	9	5	75
25.	38	23	4	3	68
26.	40	20	3	1	64
27.	40	23	18	16	97
28.	37	21	11	8	77
29.	40	21	14	10	85
30.	43	24	10	9	86

Mean = 82.06, SD = 9.42 (Cumulative Score)

Hard DDL Experimental Group – Post-test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Cumulative Score
1.	41	22	11	8	82
2.	41	22	11	7	81
3.	42	22	10	9	83
4.	40	23	20	15	98
5.	41	15	1	0	57
6.	40	21	30	27	118
7.	42	20	21	14	97
8.	43	22	3	2	70
9.	40	23	8	5	76
10.	42	26	14	7	89
11.	45	26	13	6	90
12.	38	20	2	2	62
13.	44	28	24	10	106
14.	42	25	24	13	104
15.	41	21	0	0	62
16.	40	22	19	12	93
17.	39	25	10	9	83

18.	42	22	16	10	90
19.	34	20	12	10	76
20.	44	22	8	8	82
21.	42	26	10	7	85
22.	41	24	3	2	70
23.	40	27	20	13	100
24.	38	19	12	7	76
25.	41	23	17	12	93
26.	45	21	7	4	77
27.	41	20	14	11	86
28.	42	28	7	7	84
29.	42	22	7	4	75
30.	45	28	12	10	95

Mean = 84.66, SD = 13.72 (Cumulative Score)

The following results are obtained by running an ANOVA test on the cumulative scores in SPSS:

Comparison of the Groups – Post-Test

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	9.46	2.998	.006
	Hard	30	12.06	2.998	.000
Soft	PPP	30	9.46	2.998	.006
	Hard	30	2.60	2.998	.662
Hard	PPP	30	12.06	2.998	.000
	Soft	30	2.60	2.998	.662

The above table reveals how the groups differ from each other. Examination of the P-values reveals that there is a significant difference in the scores of each group. The greatest difference lies between the Hard Experimental Group and the PPP group, which is so low that the figure provided by SPSS appears as 0.000. Moving on to the difference between the Soft Experimental Group and PPP group, the P value of

0.006 is once again very significant. These results give us more than enough grounds to reject the Null Hypothesis 1.

4.3.2.2 Null Hypothesis 2: There is no Significant Difference in the Delayed Post-Test Scores of the Three Groups.

The delayed post-test scores of the three groups (i.e. Control Group, Soft Experimental Group and Hard Experimental Group) are given below:

Control Group – Delayed Post-Test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Cumulative Score
1.	38	17	5	4	64
2.	34	18	8	6	66
3.	39	21	5	3	68
4.	40	18	11	4	73
5.	38	21	7	4	70
6.	40	19	10	9	78
7.	38	20	7	5	70
8.	41	22	4	2	69
9.	45	19	3	1	68
10.	38	20	4	3	65
11.	28	16	7	7	58
12.	39	20	4	3	66
13.	20	35	7	3	65
14.	21	12	5	2	40
15.	38	19	11	8	76
16.	35	17	5	5	62
17.	28	12	5	3	48
18.	39	22	7	5	73
19.	39	18	12	7	76
20.	22	10	3	3	38
21.	27	18	5	4	54

22.	39	16	4	4	63
23.	38	21	4	4	67
24.	36	17	5	4	62
25.	45	24	7	5	81
26.	32	17	5	1	55
27.	41	21	8	2	72
28.	21	16	1	1	39
29.	31	16	8	7	62
30.	35	17	9	7	68

Mean = 63.86, SD = 11.02 (Cumulative Score)

Soft DDL Experimental Group – Delayed Post-Test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Sum of the Score
1.	41	22	3	1	67
2.	31	19	9	7	66
3.	36	20	2	0	58
4.	42	20	11	8	81
5.	36	25	4	3	68
6.	42	19	12	9	82
7.	42	22	10	9	83
8.	43	21	4	2	70
9.	37	17	7	1	62
10.	44	21	7	4	76
11.	44	21	2	2	69
12.	40	20	11	10	81
13.	34	21	5	4	64
14.	29	20	2	1	26
15.	43	23	14	8	88
16.	41	18	10	2	71
17.	35	16	11	8	70
18.	36	23	8	5	72

19.	36	21	3	3	63
20.	32	18	2	2	54
21.	42	20	9	8	79
22.	42	22	2	1	67
23.	40	19	5	3	67
24.	43	23	6	3	75
25.	38	18	4	4	64
26.	43	23	4	4	74
27.	13	7	4	4	28
28.	42	24	5	3	74
29.	31	17	10	2	60
30.	43	20	9	9	81

Mean = 68.00, SD = 13.76 (Cumulative Score)

Hard DDL Experimental Group – Delayed Post-Test

Participant No	Essays – General DMs Score	Essays – Study DMs Score	Sentences General Score	Sentences Meanings Score	Cumulative Score
1.	42	24	10	8	84
2.	40	20	9	7	76
3.	29	16	8	7	60
4.	40	22	3	3	68
5.	34	18	3	3	58
6.	42	21	10	5	78
7.	38	17	9	3	67
8.	41	23	9	8	81
9.	43	21	18	13	95
10.	43	22	15	6	86
11.	43	24	10	8	85
12.	45	20	6	2	73
13.	43	22	19	11	95
14.	21	18	8	4	51
15.	35	22	7	4	68

16.	40	21	13	10	84
17.	38	21	13	9	81
18.	41	26	14	10	91
19.	42	23	3	1	69
20.	34	18	4	3	59
21.	41	25	4	3	73
22.	42	19	6	5	72
23.	28	15	3	2	48
24.	41	25	6	4	76
25.	37	17	2	2	58
26.	38	17	9	7	71
27.	42	22	11	8	83
28.	44	23	17	13	97
29.	42	25	13	10	90
30.	41	24	10	9	84

Mean = 75.36, SD = 13.08 (Cumulative Score)

An ANOVA test gives us the overall P value of 0.003. Once again, the P value or the significance of difference among the three groups is lower than 0.05, which means that there is a significant difference in the results of the Delayed Post-test. A Post Hoc test provides more detail about the differences between the different groups, which are tabulated below.

Comparison of the Groups – Delayed Post-Test

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	4.13	3.274	.420
	Hard	30	11.50	3.274	.002
Soft	PPP	30	4.13	3.274	.420
	Hard	30	7.367	3.274	.069
Hard	PPP	30	11.50	3.274	.002
	Soft	30	7.36	3.274	.069

It can be seen that there is still a significant difference in the results of the different groups. However, the P value for the difference between the Control group

and the Soft Experimental Group is higher than 0.05 which means that the difference is not significant. However, the overall low P value still gives us grounds to reject the null hypothesis.

4.3.2.3 Null Hypothesis 3: There is No Significant Difference in the Pre and Post-Test Scores of the Control Group.

The cumulative scores of the control group before and after treatment with the PPP method are below:

Control Group – Pre-Test versus Post-Test

Participant No	Pre-test Score	Post-test Score
1.	67	80
2.	52	84
3.	95	55
4.	60	57
5.	70	76
6.	54	66
7.	77	76
8.	32	74
9.	55	84
10.	72	80
11.	61	60
12.	70	81
13.	73	64
14.	70	60
15.	53	75
16.	76	80
17.	75	55
18.	66	88
19.	72	73
20.	76	97
21.	66	90

22.	69	76
23.	49	72
24.	54	56
25.	63	80
26.	48	77
27.	57	72
28.	67	72
29.	67	59
30.	51	59

Running a paired samples t-test on the results reveals the following statistics:

Control Group – Pre-Test versus Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	63.90	30	12.04	2.19			
Post-test	72.60	30	11.28	2.05	-2.838	29	.008
Difference	8.70						
Percentage increase	13.61%						

In the above statistics, both of the samples are similar in the distribution of the scores due to similar standard deviation values. While there is indeed a difference between the means of the two samples, the P value is a more important statistic to ascertain if the difference between the two scores is worthy of consideration or not. The P value in this case is very low, i.e. 0.008, meaning that there is a significant difference in the pre and post-test results of the control group. Hence, Null Hypothesis 3 is rejected.

4.3.2.4 Null Hypothesis 4: There is no Significant Difference in the Pre and Post-test Scores of the Soft Experimental Group.

The below table shows the scores that the learners posted in their pre-tests and post-test:

Soft Experimental Group – Pre-Test versus Post-Test

Participant No	Pre-test Score	Post-test Score
1.	77	74
2.	70	94
3.	74	80
4.	64	97
5.	77	96
6.	30	81
7.	78	93
8.	65	77
9.	77	73
10.	60	61
11.	66	82
12.	80	80
13.	71	93
14.	70	74
15.	35	85
16.	59	80
17.	54	83
18.	64	77
19.	68	93
20.	65	85
21.	29	90
22.	81	83
23.	68	79
24.	85	75
25.	72	68
26.	69	64
27.	66	97
28.	66	77
29.	53	85
30.	66	86

A paired samples t-test is used to find out the difference between the pre-test and post-test scores of the same group. The results of such a post-test for the current group are found below:

Control Group – Pre-Test versus Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	65.30	30	13.74936	2.51028			
Post-test	82.06	30	9.42825	1.72136	-5.257	29	.000
Difference	16.76667						
Percentage increase	25.16%						

The above tables show us the difference in the means of the two groups. Besides having visibly different means, the t-test reveals that the difference is extremely significant; therefore Null Hypothesis 4 is rejected.

4.3.2.5 Null Hypothesis 5: There is No Significant Difference in the Pre and Post-Test Scores of the Hard Experimental Group.

The scores of the participants of the Hard Experimental group in their pre and post-tests are found in the below tables:

Hard Experimental Group – Pre-Test versus Post-Test

Participant No	Pre-test Score	Post-test Score
1.	60	82
2.	77	81
3.	45	83
4.	63	98
5.	67	57
6.	74	118
7.	63	97
8.	50	70
9.	65	76
10.	83	89
11.	70	90
12.	68	62
13.	65	106
14.	65	104
15.	82	62
16.	59	93
17.	69	83
18.	74	90
19.	55	76
20.	35	82
21.	68	85
22.	66	70
23.	72	100
24.	88	76
25.	64	93
26.	67	77
27.	60	86
28.	68	84
29.	54	75
30.	73	95

Running a paired t-test reveals the following statistics:

Control Group – Pre-Test versus Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	65.6333	30	10.93392	1.99625			
Post-test	84.6667	30	13.72472	2.50578	-6.134	29	.000
Difference	19.03						
Percentage increase	28.99%						

The above statistics produced by SPSS indicate that the P value of is really low (i.e. 0.000). This means that it is too low to show up in the first three values after the decimal. This is lower than 0.05 which means that the change in the scores is very significant. In addition, it can be observed that the change in means of the scores between the pre and post tests is the highest as compared to other groups. Null hypothesis 6 is, therefore, rejected.

4.3.2.6 Null Hypothesis 6: There is No Significant Difference in the Post-test and Delayed Post-Test Scores of the Control Group.

The post and delayed post-test scores of the control group are the following:

Control Group – Post-Test versus Delayed Post-Test

Participant No	Post-test Score	Delayed Post-test Score
1.	80	64
2.	84	66
3.	55	68
4.	53	73
5.	76	70
6.	65	78
7.	76	70
8.	74	69
9.	84	68

10.	80	65
11.	56	58
12.	81	66
13.	64	65
14.	60	40
15.	75	76
16.	80	62
17.	55	48
18.	88	73
19.	71	76
20.	97	38
21.	90	54
22.	62	63
23.	72	67
24.	51	62
25.	80	81
26.	77	55
27.	72	72
28.	72	39
29.	48	62
30.	60	68

A paired t-test reveals the following results:

Control Group – Post-Test versus Delayed Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	71.2667	30	12.59155	2.29889			
Post-test	63.8667	30	11.02578	2.01302	2.401	29	.023
Difference	7.40						
Percentage increase	11.26%						

The results show that there is a significant change in the scores between the post-test and delayed post-test of the control group (i.e. P = 0.023) and that the means decreased considerably in the delayed post-tests, hence the null hypothesis is rejected.

4.3.2.7 Null Hypothesis 7: There is No Significant Difference in the Post-Test and Delayed Post-Test Scores of the Soft Experimental Group.

The post-test and delayed post-test of the Soft Experimental Group are tabulated below:

Soft Experimental Group

Participant No	Post-test Score	Delayed Post-test Score
1.	74	67
2.	94	66
3.	80	58
4.	97	81
5.	96	68
6.	81	82
7.	93	83
8.	77	70
9.	73	62
10.	61	76
11.	67	69
12.	80	81
13.	93	64
14.	26	26
15.	85	88
16.	68	71
17.	79	70
18.	77	72
19.	93	63
20.	61	54
21.	90	79

22.	83	67
23.	79	67
24.	75	75
25.	68	64
26.	64	74
27.	42	28
28.	77	74
29.	85	60
30.	82	81

Control Group

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	76.6667	30	15.55044	2.83911			
Post-test	68.0000	30	13.76152	2.51250	4.049	29	.000
Difference	8.66						
Percentage increase	10.52%						

The statistical test shows that the difference in the means is significant with the P value being 0.000, giving us grounds to reject the null hypothesis.

4.3.2.8 Null Hypothesis 8: There is No Significant Difference in the Post-Test and Delayed Post-Test Scores of the Hard Experimental Group.

The scores of the participants of the Hard Experimental Group in their post-test and delayed post-test scores are the following:

Hard Experimental Group – Post-Test versus Delayed Post-Test

Participant No	Post-test Score	Delayed Post-test Score
1.	82	84
2.	81	76
3.	83	60
4.	98	68
5.	57	58
6.	118	78
7.	97	67
8.	70	81
9.	76	95
10.	89	86
11.	90	85
12.	62	73
13.	106	95
14.	104	51
15.	62	68
16.	93	84
17.	83	81
18.	90	91
19.	76	69
20.	82	59
21.	85	73
22.	70	72
23.	100	48
24.	76	76
25.	93	58
26.	77	71
27.	86	83
28.	84	97
29.	75	90
30.	95	84

Control Group – Post-Test versus Delayed Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pre-test	84.6667	30	13.72472	84.6667			
Post-test	75.3667	30	13.08720	75.3667	2.707	29	.011
Difference	9.30						
Percentage increase	10.71%						

The above two tables show us the statistics for the differences in the post-test and delayed post-test scores for the Hard DDL Experimental group. It can be observed that the significance value is 0.011, which is less than the threshold of 0.05. This means that significant changes in scores occurred between the two tests, thus, Null Hypothesis 8 is also rejected.

4.3.3 Comparison of HEG and Control Group in Post-Tests

In this section, the results of the experimental group and the control group. The control group here was the group of learners taught using the traditional PPP method. The HEG was taught using the hard version of Data-Driven Learning Method, i.e. using corpus software on a computer. It is hypothesized in Hypothesis No. 1 that the Hard Experimental group yields better results than the traditional PPP classroom method in post-tests. To this end, the post-test scores of the two groups must be compared. Since it has already been found in Null Hypothesis 1 that the two groups were comparable in the pre-tests, i.e. the similar scores mean that on average the learners were on the same level before starting the treatment, the post-test scores of the two groups can now be compared to see if any difference in the scores occurred.

The following is a detailed comparison of the post-test scores of the HEG and the SEG.

4.3.3.1 Cumulative Scores.

The cumulative scores of the two groups are tabulated below.

Cumulative Scores – Post-Test

Participant No	Control Group	Hard Experimental Group
1.	80	82
2.	84	81
3.	55	83
4.	57	98
5.	76	57
6.	66	118
7.	76	97
8.	74	70
9.	84	76
10.	80	89
11.	60	90
12.	81	62
13.	64	106
14.	60	104
15.	75	62
16.	80	93
17.	55	83
18.	88	90
19.	73	76
20.	97	82
21.	90	85
22.	76	70
23.	72	100
24.	56	76
25.	80	93
26.	77	77
27.	72	86

28.	72	84
29.	59	75
30.	59	95

The test reveals the following statistics:

Cumulative Scores – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)	Improvement
Control	30	72.60	11.282	2.060				13.61%
HEG	30	84.67	13.725	2.506	-3.720	58	.000	28.99%
Difference	12.06							

The first and most important result here is the significance of the results. SPSS shows the P value as 0.000 which means it is really low. This means that the significance of these results is high, indicating that the treatment of the hard experimental group has been more effective than the control group. An interesting thing to note is that the standard deviation values for both groups, i.e. 13.26 and 10.93 for the control and hard experimental groups respectively in the pre-test and 11.28 and 13.72 for the control and hard experimental groups respectively in the post-test. This shows that the PPP method reduced the overall variance in the learners and learners were able to close the gap somewhat between them. In contrast, the standard deviation in the hard experimental group increased noticeably; this means that the variance increased, indicating that the scores for some learners increased more than the others.

Coming to the improvement in scores, the mean of the scores of the control group showed an improvement of 13.61% while there was an increase of 28.99% in the mean of scores of the hard experimental group. The improvement in the scores of the hard experimental group is more than double of the control group. The above statistics provide us our first look into the overall trends found in the study, which seem to favor the HEG. This provides the an indicator towards the merits of the factors that are believed to form the basis of DDL, such as consciousness raising, noticing, increased involvement load and activation of cognitive skills.

To further explore the intricacies of the differences between the two groups, deeper analysis into the components of the data is required, which is performed below.

4.3.3.2 Sentence Making Exercise.

The participants of both of the groups had the following scores in the post-tests:

Sentence Making Exercise: General Score – Post-Test

Participant No	Control Group	Hard Experimental Group
1.	37	41
2.	42	41
3.	37	42
4.	33	40
5.	43	41
6.	33	40
7.	45	42
8.	42	43
9.	39	40
10.	37	42
11.	30	45
12.	40	38
13.	35	44
14.	39	42
15.	37	41
16.	40	40
17.	21	39
18.	40	42
19.	40	34
20.	40	44
21.	39	42
22.	44	41

23.	38	40
24.	31	38
25.	41	41
26.	39	45
27.	38	41
28.	40	42
29.	31	42
30.	35	45

Sentence Making Exercise: General Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	30	37.53	4.897	.894			
Hard Experimental Group	30	41.27	2.288	.418	-3.783	58	.000
Difference	8.44						
Percentage difference	9.49%						

The t-test shows that there is a significant difference in the general scores of the sentence-making exercise between the two groups with the P value presenting as 0.000. The hard experimental group showed higher scores than the control group by 9.49%. We can also see that the standard deviation in the hard experimental group is lower than the control group. This means that the scores of the learners in the hard experimental group were generally closer to each other, showing more consistent improvement. Another thing to explore here is if the learners were able to diversify their use of the DMs, i.e. if they used the DMs in more than one way. The Sentences Meanings Score are given below for this purpose:

Sentence Making Exercise: Meanings Score – Post-Test

Participant No	Control Group	Hard Experimental Group
1.	22	22
2.	20	22
3.	14	22
4.	16	23
5.	28	15
6.	15	21
7.	22	20
8.	20	22
9.	19	23
10.	23	26
11.	19	26
12.	22	20
13.	21	28
14.	21	25
15.	22	21
16.	24	22
17.	17	25
18.	22	22
19.	17	20
20.	25	22
21.	22	26
22.	17	24
23.	20	27
24.	17	19
25.	21	23
26.	20	21
27.	21	20
28.	22	28
29.	17	22
30.	16	28

An independent samples t-test gives the following results:

Sentence Making Exercise: Meanings Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	20.07	30	3.140	.573	-3.486	58	.001
Hard Experimental Group	22.83	30	3.007	.549			
Difference	2.76						
Percentage difference	12.86%						

The P value can be observed to be 0.001 which is lower than the threshold of 0.05. This implies that the difference between the two groups is significant. The results mean that the learners in the hard experimental group showed more improvement in the different ways they use sentences in comparison with the PPP group. A notable thing here is the difference in the means of the two groups for the meanings score, which is 12.86% while the percentage difference in the means of the general score in sentence making is 9.49%. This means that the hard experimental group provided greater improvement in how diversely the learners use DMs than just simple improvement in the ability to use them.

4.3.3.3 Essays.

Similar to the sentence making exercises, there are also two essays scores: the General DMs score and the Study DMs score. The general DMs score represents the overall number of DMs used in a sample while the Study DMs score represents the use of the DMs that were taught in this research.

The general DMs score is discussed first and the scores are given below:

Essays: General DMs Score – Post-Test

Participant No	Control Group	Hard Experimental Group
1.	12	11
2.	14	11
3.	2	10
4.	4	20
5.	2	1
6.	10	30
7.	5	21
8.	6	3
9.	16	8
10.	12	14
11.	6	13
12.	16	2
13.	5	24
14.	0	24
15.	10	0
16.	9	19
17.	10	10
18.	16	16
19.	8	12
20.	23	8
21.	17	10
22.	8	3
23.	9	20
24.	4	12
25.	12	17
26.	12	7
27.	9	14
28.	6	7
29.	6	7
30.	4	12

An independent samples t-test in SPSS yields the following results:

Essays: General DMs Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	9.10	30	5.235	.956			
Hard Experimental Group	12.20	30	7.350	1.342	-1.882	58	.065
Difference	3.10						
Percentage difference	29.10%						

The results become interesting for this component of the test, because the P value is 0.065, which is over the 0.05 threshold for the significance of the results. This means that the difference in the score can be due to chance. This means that it is possible that the two groups had similar improvements in how many general DMs they used on average. However, this does not mean that there the results shall be completely disregarded, as the P value is not hugely above the 0.05 threshold and there could still be some merit to these results. In percentage terms, the learners in the hard experimental group scored 29.10% more than the control group. In addition, there is an extra bit of context to be found in the results of the study DMs score. This score counts the number of study DMs used by each learner.

Essays: Study DMs Score – Post-Test

Participant No	Control Group	Hard Experimental Group
1.	9	8
2.	8	7
3.	2	9
4.	4	15
5.	3	0
6.	8	27
7.	4	14
8.	6	2

9.	10	5
10.	8	7
11.	5	6
12.	3	2
13.	3	10
14.	0	13
15.	6	0
16.	7	12
17.	7	9
18.	10	10
19.	8	10
20.	9	8
21.	12	7
22.	7	2
23.	5	13
24.	4	7
25.	6	12
26.	6	4
27.	4	11
28.	4	7
29.	5	4
30.	4	10

Running an independent samples t-test yields the following results:

Essays: Study DMs Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	5.90	30	2.695	.492			
Hard Experimental Group	8.37	30	5.372	.981	-2.248	58	.028
Difference	2.46						
Percentage difference	34.61%						

It can be observed that the P value for the study-specific DMs scores is lower than the threshold, meaning that the difference between the two groups is significant. The percentage difference between the two means is 34.61%. This means that while there wasn't a significant difference in the general DM score, the participants of the hard experimental group demonstrated increased usage of the DMs taught in the study. Another important statistic here is to look at how much the study-specific DMs comprised all of the DMs in the pre and post-tests. For the control group, in the pre-test, 63.5% of the DMs were the ones taught in the study, while 64.8% DMs were the study-specific DMs in the post-test. For the hard experimental group, in the pre-test, 60.7% of the DMs were the ones taught in the study, while 68.5% DMs were the study-specific DMs in the post-test. This shows that the learners in the hard experimental group showed greater improvement in the study-specific DMs than the control group.

4.3.3.4 Discussion

The comparison of the two groups showed that the Hard DDL method had significantly higher outcomes for the participants of its respective group in comparison with the PPP method. The higher improvement was consistently found in all the different components of the data. The results reflect the components of the theoretical framework of the research. Hard DDL encapsulates all the features of the Involvement Load Hypothesis by Hulstijn and Laufer (2001) as well as Schmidt's (1990) Noticing Hypothesis. The process of learning using Hard DDL involves a process of search, which is motivated by the need to understand and use the language. The process of evaluation is performed in order to decipher the underlying rules and add them to the learner's arsenal. The questions of the instructor as well as the ability of the corpus to highlight the keyword under study also increase the noticing. The interactive and computer enabled nature of Hard DDL is also more cognitively stimulating, leading the learners to develop better understanding. It is directly in line with the beliefs held by Constructivism, which considers language learning to be a dynamic process, dependent on problem solving and meaning making. Hard DDL comprises elements that stimulate the cognitive faculties of a person, leading to increased consciousness of language patterns.

The above results are consistent with the findings of Bao (2021), who conducted analysis of 79 research articles and reported that DDL has generally positive effects in improving learner proficiency. Similar results can also be seen in studies conducted by Barabadi and Khajavi (2017), Boontam and Phoocharoensil (2018) and Shah (2021). The participants of the Hard DDL outperformed the learners administered with the PPP method in every component. They showed higher scores in not only the general use of DMs in the sentences, but also demonstrated the ability to use them in more diverse ways. It was also found that the learners used more DMs in general in comparison with the control group, meaning that Hard DDL brings attention to the use of DMs overall.

4.3.4 Comparison of the SEG and Control Group

The second hypothesis of this study deals with the difference between the second experimental group, i.e., the soft experimental group, and the control group. The control group is once again the group of learners taught using the traditional PPP method. The Soft Experimental group comprised the learners treated with the soft version of the DDL method, i.e. using printed handouts with examples of the DMs picked from a corpus. The hypothesis postulates that the participants of the Soft Experimental group yield better results than the participants treated with the traditional PPP classroom method. A comparison between the two groups will be needed to either substantiate or reject this hypothesis. It has already been established earlier that the two groups were comparable in the pre-test and hence are appropriate for this study.

4.3.4.1 Cumulative Scores.

Testing the Null Hypothesis 1 demonstrated that there is a significant difference between the post-test results of the two groups. For further analysis, the cumulative scores of the two groups are below.

Cumulative Scores – Post-Test

Participant No	Control Group	Soft Experimental Group
1.	80	74
2.	84	94
3.	55	80
4.	57	97
5.	76	96
6.	66	81
7.	76	93
8.	74	77
9.	84	73
10.	80	61
11.	60	82
12.	81	80
13.	64	93
14.	60	74
15.	75	85
16.	80	80
17.	55	83
18.	88	77
19.	73	93
20.	97	85
21.	90	90
22.	76	83
23.	72	79
24.	56	75
25.	80	68
26.	77	64
27.	72	97
28.	72	77
29.	59	85
30.	59	86

The test reveals the following statistics:

Cumulative Scores

Group	Mean	N	Std. Deviation	Std. Error Mean	df	Sig. (2-tailed)
Control Group	72.60	30	11.282	2.060	58	.001
Hard Experimental Group	82.07	30	9.428	1.721		
Difference	9.46					
Percentage difference	12.24%					

Coming to the improvement in scores, the mean of the scores of the control group showed an improvement of 13.61% while the soft experimental group showed an increase of 22.75%. This difference in the improvement in the scores of the two groups is significant and shows that the treatment administered to the soft experimental was more effective.

4.3.4.2 Sentence Making Exercise.

Coming to the individual components of the tests, the general scores of the learners in the Sentence Making Exercise are below.

Sentence Making Exercise – General Score

Participant No	Control Group	Soft Experimental Group
1.	37	41
2.	42	42
3.	37	39
4.	33	44
5.	43	40
6.	33	44
7.	45	40

8.	42	42
9.	39	37
10.	37	42
11.	30	41
12.	40	40
13.	35	40
14.	39	37
15.	37	44
16.	40	40
17.	21	42
18.	40	40
19.	40	41
20.	40	42
21.	39	45
22.	44	41
23.	38	37
24.	31	40
25.	41	38
26.	39	40
27.	38	40
28.	40	37
29.	31	40
30.	35	43

Sentence Making Exercise – General Score

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	37.53	30	4.897	.894			
Hard Experimental Group	40.63	30	2.157	.394	-3.173	58	.002
Difference	3.1						
Percentage difference	7.93%						

The results of the t-test indicate a P value of 0.002, which shows that the difference in the general scores of the sentence-making exercise of the two groups is significant and not an occurrence by chance. The participants of the soft experimental group scored 7.93% more than the participants of the control group. The lower standard deviation also indicates that the scores of the learners in the soft experimental group were less spread out, meaning that there was more consistent improvement.

The Sentences Meanings scores provide further insight into how the DMs were used in making sentences, i.e. how diversely they were used in the sentences.

Sentence Making Exercise: Meanings Score – Post-Test

Participant No	Control Group	Soft Experimental Group
1.	22	21
2.	20	22
3.	14	20
4.	16	25
5.	28	24
6.	15	25
7.	22	21
8.	20	22
9.	19	22
10.	23	2
11.	19	25
12.	22	17
13.	21	20
14.	21	18
15.	22	26
16.	24	21
17.	17	22
18.	22	24
19.	17	21
20.	25	21

21.	22	24
22.	17	19
23.	20	22
24.	17	21
25.	21	23
26.	20	20
27.	21	23
28.	22	21
29.	17	21
30.	16	24

The following table shows the results of the analysis of these values using an independent samples t-test in SPSS.

Sentence Making Exercise: Meanings Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	20.07	30	3.140	.573			
Hard Experimental Group	21.23	30	4.207	.768	-1.217	58	.228
Difference	1.16						
Percentage difference	5.61%						

The P value for this pair of scores is 0.228 which is significantly higher than 0.05. This means that the difference found in the scores is not statistically significant. To elaborate, this means that the difference seen in the scores of the two groups could be coincidental or by chance, and therefore this difference cannot be relied on to infer any results. The results show that the difference in the improvement of scores is negligible at 5.61% and that both groups posted similar improvement, i.e. the participants of both groups used the DMs in sentences in similarly different ways.

4.3.4.3 Essays.

The General DMs scores in essays for the both groups are given below.

Essays: General DMs Score – Post-Test

Participant No	Control Group	Soft Experimental Group
1.	12	6
2.	14	18
3.	2	14
4.	4	16
5.	2	21
6.	10	7
7.	5	19
8.	6	9
9.	16	9
10.	12	11
11.	6	8
12.	16	14
13.	5	21
14.	0	11
15.	10	8
16.	9	10
17.	10	11
18.	16	8
19.	8	20
20.	23	12
21.	17	11
22.	8	15
23.	9	12
24.	4	9
25.	12	4
26.	12	3
27.	9	18

28.	6	11
29.	6	14
30.	4	10

An independent samples t-test in SPSS provides the following results:

Essays –General DMs Score

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	9.10	30	5.235	.956	-2.229	58	.030
Hard Experimental Group	12.00	30	4.835	.883			
Difference	2.9						
Percentage difference	27.48%						

In contrast with the previous section of the results, the P value for this score is 0.030, which is lower than 0.05. This indicates that the difference in the scores of the two groups is significant and not a chance occurrence. The participants of the soft experimental group showed higher levels of improvement in their scores as compared to the control group, with their scores being 27.48% higher than the control group. This shows that the treatment in the soft experimental group was able to provide significantly more improvement than the control group.

The scores of study-specific DMs for the two groups can provide further insight.

Essays: Study DMs Score – Post-Test

Participant No	Control Group	Soft Experimental Group
1.	9	6
2.	8	12
3.	2	7
4.	4	12
5.	3	11
6.	8	5
7.	4	13
8.	6	4
9.	10	5
10.	8	6
11.	5	8
12.	3	9
13.	3	12
14.	0	8
15.	6	7
16.	7	9
17.	7	8
18.	10	5
19.	8	11
20.	9	10
21.	12	10
22.	7	8
23.	5	8
24.	4	5
25.	6	3
26.	6	1
27.	4	16
28.	4	8
29.	5	10
30.	4	9

Running an independent samples t-test yields the following results:

Essays: Study DMs Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	5.90	30	2.695	.492	-2.988	58	.004
Hard Experimental Group	8.20	30	3.242	.592			
Difference	2.3						
Percentage difference	32.62%						

For the study-specific DMs, the statistical test again shows a very significant P value of 0.004, which implies that the difference between the two groups is very significant. The percentage difference between the two means is 32.62%. The participants of the soft experimental group demonstrated increased usage of the DMs that were taught in the study in comparison with the control group.

An interesting thing to look at here is how much of the total score the study-specific DMs comprised in the pre and post-tests. In the pre-test, the study-specific DMs made up 63.5% of the total DM use by the participants of the control group, while this percentage was 52.7% for the soft experimental group. In the post-test, this percentage for the control group increased to 64.8% and to 68.63% for the soft experimental group. Thus, the ratio of the study DMs among overall DMs for the soft experimental group climbed past the control group in the post-tests and showed great improvement.

4.3.4.4 Discussion.

The results indicate that Soft DDL group provided higher improvement for the learners in comparison with the control group. This is consistent with the scores of the hard DDL, with the exception of the meanings score in the sentence making, where the P value for the difference was too high for any difference to be considered

statistically significant. This is against the trend in the rest of the scores which were significantly higher than the control group. An explanation for this stems from the theoretical framework of this research, which focuses on noticing (Schmidt, 1990) and the need, search and evaluation components of the involvement load hypothesis (Hulstijn and Laufer, 2001). It appears that the soft version of DDL did not sufficiently engage the higher-order cognitive skills of the learners. It could further be theorized that the soft form was not equally stimulating as the computer based version in order for it to increase noticing and involvement load, so that it would provide increased results in all kinds of activities. However, it must still be noted that the results from the other components are significant enough to substantiate the belief that soft DDL has better outcomes for the learners.

The results of this study echo the findings reported by other researchers such as Huang (2014), Klomkaew and Boontam (2023) and Smart (2014), who have also found the paper-based version of DDL to be an effective method of teaching various grammatical and linguistic items ranging from passive voice to linguistic patterns to conditional sentences. This aligns with the general trend found in employing DDL in classrooms and demonstrates that while there are indeed differences in the results between the two versions, they are consistent with each other in performing better than the control group.

4.3.5 Comparison of SEG and HEG in Post-Tests

The third hypothesis is about the comparison between the two experimental groups. The hypothesis postulates that the difference between the two groups should be negligible. This section explores if this hypothesis can be held true based on the data obtained during the study.

While it has already been established in the discussion of Null Hypothesis 1 that the difference in the cumulative scores of the two groups is not significant, more statistical tests are required to provide a clearer picture of the differences in the scores.

4.3.5.1 Cumulative Scores.

Cumulative Scores – Post-test

Participant No	Hard Experimental Group	Soft Experimental Group
1.	82	74
2.	81	94
3.	83	80
4.	98	97
5.	57	96
6.	118	81
7.	97	93
8.	70	77
9.	76	73
10.	89	61
11.	90	82
12.	62	80
13.	106	93
14.	104	74
15.	62	85
16.	93	80
17.	83	83
18.	90	77
19.	76	93
20.	82	85
21.	85	90
22.	70	83
23.	100	79
24.	76	75
25.	93	68
26.	77	64
27.	86	97
28.	84	77
29.	75	85

30.	95	86

A t-test reveals the following statistics:

Cumulative Scores – Post-test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	82.07	30	9.428	1.721	-.855	58	.396
Hard Experimental Group	84.67	30	13.725	2.506			
Difference	2.6						
Percentage difference	3.11%						

The above results give us a P value of 0.396 which is significantly higher than the threshold of 0.05. This means that any observed difference in the values may have occurred due to chance. If we even look at the percentage difference notwithstanding the significance value, the value comes out to 3.11% which is very little, therefore providing grounds to believe that the scores posted by the participants of the two groups are similar. Examination of the difference in the standard deviation, it can be seen that it is lower for the soft experimental group compared to the hard experimental group. This means that the participants of the latter group started to equalize more in the post-test in comparison with the hard experimental group.

4.3.5.2 Sentence Making Exercise.

The individual tests provide more data to determine if there are any significant differences to be found in any of them. The first under discussion is the general score of the learners in the Sentence Making Exercise.

Sentence Making Exercise: General Score – Post-Test

Participant No	Hard Experimental Group	Soft Experimental Group
1.	41	41
2.	41	42
3.	42	39
4.	40	44
5.	41	40
6.	40	44
7.	42	40
8.	43	42
9.	40	37
10.	42	42
11.	45	41
12.	38	40
13.	44	40
14.	42	37
15.	41	44
16.	40	40
17.	39	42
18.	42	40
19.	34	41
20.	44	42
21.	42	45
22.	41	41
23.	40	37
24.	38	40
25.	41	38
26.	45	40
27.	41	40
28.	42	37
29.	42	40
30.	45	43

An independent samples t-test on the above figures provides the following statistics:

Sentence Making Exercise: General Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	40.63	30	2.157	.394			
Hard Experimental Group	41.27	30	2.288	.418	-1.103	58	.275
Difference	0.63						
Percentage difference	1.56%						

Consistent with previous results, this statistical test gives us a P value of 0.275, which indicates that the difference between the two groups is negligible. Furthermore, it shall be noted that the standard deviation for the two groups is similar, which means that the scores were similar across the two groups for the participants.

Moving on to the meaning scores, the participants had the following scores for the diversity in meanings in sentence making exercise:

Sentence Making Exercise: Meanings Score – Post-Test

Participant No	Hard Experimental Group	Soft Experimental Group
1.	22	21
2.	22	22
3.	22	20
4.	23	25
5.	15	24
6.	21	25
7.	20	21
8.	22	22
9.	23	22

10.	26	2
11.	26	25
12.	20	17
13.	28	20
14.	25	18
15.	21	26
16.	22	21
17.	25	22
18.	22	24
19.	20	21
20.	22	21
21.	26	24
22.	24	19
23.	27	22
24.	19	21
25.	23	23
26.	21	20
27.	20	23
28.	28	21
29.	22	21
30.	28	24

Applying an independent samples t-test in SPSS to analyze the data provides the following results:

Sentence Making Exercise: Meanings Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	21.23	30	4.20	.768			
Hard Experimental Group	22.83	30	3.00	.549	-1.695	58	.096
Difference	1.6						
Percentage difference	33.33%						

The significance value again is higher than the threshold of 0.05, being produced as 0.096. This means that the difference found in the scores is not statistically significant to be considered valid for inferring, i.e. the scores posted by the two groups are similar. The standard deviation is also similar for the two groups.

4.3.5.3 Essays.

The essays general DMs scores for the two groups are below.

Essays: General DMs Score – Post-Test

Participant No	Hard Experimental Group	Soft Experimental Group
1.	11	6
2.	11	18
3.	10	14
4.	20	16
5.	1	21
6.	30	7
7.	21	19
8.	3	9
9.	8	9
10.	14	11
11.	13	8
12.	2	14
13.	24	21
14.	24	11
15.	0	8
16.	19	10
17.	10	11
18.	16	8
19.	12	20
20.	8	12
21.	10	11
22.	3	15

23.	20	12
24.	12	9
25.	17	4
26.	7	3
27.	14	18
28.	7	11
29.	7	14
30.	12	10

An independent samples t-test in SPSS provides the following results:

Sentence Making Exercise: Meanings Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	12.00	30	4.835	.883	-.125	58	.901
Hard Experimental Group	12.20	30	7.350	1.342			
Difference	0.20						
Percentage difference	1.65%						

The significance value for this score is once again very high sitting P value for this score is 0.90, which is significantly higher than 0.05, indicating that the difference between the two groups is not significant.

Finally, the scores as well as the statistics for the differences in the use of study-specific DMs between the two groups are provided below.

Essays: Study DMs Score – Post-Test

Participant No	Hard Experimental Group	Soft Experimental Group
1.	8	6
2.	7	12
3.	9	7
4.	15	12
5.	0	11
6.	27	5
7.	14	13
8.	2	4
9.	5	5
10.	7	6
11.	6	8
12.	2	9
13.	10	12
14.	13	8
15.	0	7
16.	12	9
17.	9	8
18.	10	5
19.	10	11
20.	8	10
21.	7	10
22.	2	8
23.	13	8
24.	7	5
25.	12	3
26.	4	1
27.	11	16
28.	7	8
29.	4	10
30.	10	9

An independent samples t-test in SPSS yields the following results:

Essays: Study DMs Score – Post-Test

Group	Mean	N	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Control Group	8.20	30	3.242	.592	-.145	58	.885
Hard Experimental Group	8.37	30	5.372	.981			
Difference	0.16						
Percentage difference	2.05%						

For the study-specific DMs, the statistical test again gives us a very low P value of 0.885, once again meaning that there is no significant difference between the two groups.

The above discussion provides enough substance to accept the hypothesis that the two experimental groups yielded similar results.

4.3.5.3 Discussion.

The results reveal that the post-test results of the hard and soft DDL groups were similar. Notably, the P values for the differences between all the tests returned higher than the threshold, meaning that any difference in the results could be merely coincidental. There is very little empirical work on comparing the two kinds of DDL, as most of the work comparing the two versions is theoretical (Boulton, 2009). The only notable study is from Chujo et al. (2012) who worked with Japanese learners and reported that the paper-based soft version of DDL provided similar results to the computer-based hard version.

Despite the lower P values, it shall be noted that the scores of the hard DDL group are consistently higher than the soft DDL. If these results were to be considered, it can be hypothesized that any difference in the results may be occurring due to the different natures of the interaction of the two types of DDL methodologies.

The theoretical basis of DDL relies on concepts such as consciousness raising, noticing of the language concepts, application of cognitive skills on them to make meaning and involvement load, all of which could affect how successful a learner is in acquiring a certain linguistic item. On comparison between the two versions of the DDL, it is notable that the hard version of DDL allows the learners to actively interact with the corpus, provides a higher number of instances for the learners to look from and enables them to access broader contexts of the DMs. All of these features lead to increased learner interest in the activities, resulting in increased involvement load and higher levels of noticing as well as cognitive stimulation, which aid language learning according to Schmidt (1990) and Hulstijn and Laufer (2001).

4.3.6 Comparison of the Delayed Post-Test Scores

This hypothesis deals with comparing the delayed post-test results of the two experimental groups with the control group. It proposes that both of the versions of DDL allow better long-term retention of the DMs and their usage. The hypothesis is tested in the following sections.

4.3.6.1 Cumulative Scores.

The cumulative scores of all three groups in the delayed post-test are the following:

Cumulative Scores – Delayed Post-Tests

Participant No	Control Group	Soft Experimental Group	Hard Experimental Group
1.	64	67	84
2.	66	66	76
3.	68	58	60
4.	73	81	68
5.	70	68	58
6.	78	82	78
7.	70	83	67
8.	69	70	81
9.	68	62	95

10.	65	76	86
11.	58	69	85
12.	66	81	73
13.	65	64	95
14.	40	26	51
15.	76	88	68
16.	62	71	84
17.	48	70	81
18.	73	72	91
19.	76	63	69
20.	38	54	59
21.	54	79	73
22.	63	67	72
23.	67	67	48
24.	62	75	76
25.	81	64	58
26.	55	74	71
27.	72	28	83
28.	39	74	97
29.	62	60	90
30.	68	81	84
Mean	63.86	68	75.36
SD	11.026	13.762	13.087

The Null Hypothesis 2 has already established three main things: firstly, the difference between the scores of the control group and the soft experimental group is not significant with the P value of 0.420, which is higher than the threshold value of 0.05; secondly, there is significant difference between the hard experimental group and the control group with a P value of 0.002; and lastly, the P value for the difference between the hard and soft experimental groups is 0.069, making it an insignificant difference. The results provided by the ANOVA test are presented in the below table.

Comparison of the Three Groups – Delayed Post-Tests

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	-4.133	3.274	.420
	Hard	30	-11.500	3.274	.002
Soft	PPP	30	4.133	3.274	.420
	Hard	30	-7.367	3.274	.069
Hard	PPP	30	11.500	3.274	.002
	Soft	30	7.367	3.274	.069

The scores of the participants of the hard experimental group are 16.5% higher than the control group and 10.2% higher than the soft experimental group. Meanwhile, the results of the soft experimental group show a difference of 6.27% from the control group; however, the high P value indicates that these results are not significant.

If the change in scores between the post-test and the delayed post-test is compared, the scores of the control group decreased by 12.8%, the scores of the soft experimental group decreased by 18.73% and the results of the hard experimental group decreased by 11.62%. This indicates that the hard experimental group was overall the best in long-term retention of the concepts. While the soft experimental group showed significantly better scores in the post-test in comparison with the control group, the scores went down by a lot and somewhat became closer to the control group.

For the overall standing of the scores in comparison with the pre-test, the scores of the control group remained stagnated with the mean scores actually being 3.1% lower in the delayed post-test than the pre-test. This was, however, different in cases of the hard experimental group and the soft experimental groups, with hard experimental group showing 13.80% higher scores over long term and the soft experimental group showing 4.05% improvement.

4.3.6.2 Sentence Making Exercise.

Digging deeper, it is now time to examine the change in scores of the components of the tests. The general sentences score shall be discussed first.

Sentence Making Exercise: General Score – Delayed Post-Tests

Participant No	Control Group	Soft Experimental Group	Hard Experimental Group
1.	38	41	42
2.	34	31	40
3.	39	36	29
4.	40	42	40
5.	38	36	34
6.	40	42	42
7.	38	42	38
8.	41	43	41
9.	45	37	43
10.	38	44	43
11.	28	44	43
12.	39	40	45
13.	20	34	43
14.	21	29	21
15.	38	43	35
16.	35	41	40
17.	28	35	38
18.	39	36	41
19.	39	36	42
20.	22	32	34
21.	27	42	41
22.	39	42	42
23.	38	40	28
24.	36	43	41
25.	45	38	37
26.	32	43	38

27.	41	13	42
28.	21	42	44
29.	31	31	42
30.	35	43	41
Mean	34.83	38.03	39
SD	7.023	6.43	5.31

Running an ANOVA test on the scores gives the following results:

Comparison of the Three Groups – Delayed Post-Tests

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	-3.200	1.626	.126
	Hard	30	-4.167	1.626	.032
Soft	PPP	30	3.200	1.626	.126
	Hard	30	-.967	1.626	.823
Hard	PPP	30	4.167	1.626	.032
	Soft	30	.967	1.626	.823

This table shows how the groups differ from each other. The significance value for the difference between the control and the soft experimental groups is once again very high, sitting at 0.126, indicating that the difference is not significant and therefore the difference in the scores cannot be considered to be a reliable result to be used for interpretation. The P value for the difference between the control and the hard experimental group, however, is lower than 0.05, which means that it is a significant result. This substantiates earlier results which showed that the hard experimental group allows better retention of the DMs than the control group. Finally, coming to the difference between the hard and soft experimental groups, it can be observed that the P value is very high, meaning that the difference between the two groups is negligible. This brings up an interesting point, however, because the difference between the hard and soft experimental groups is not enough to consider the result significant; however, the result of one is significantly different than the control group, but the result of the other one is not. It seems that the scores of the soft experimental group lie between the hard experimental group and the control group and therefore behave in this way.

Coming to the percentage differences, the scores of the participants of the hard experimental group are 11.29% more than the control group and 2.51% more than the soft experimental group. Meanwhile, the results of the soft experimental group show a difference of 8.78% with the control group; however, the high P value indicates that these results are not significant.

Moving on, the sentences meaning scores in the delayed post-tests are tabulated below:

Sentence Making Exercise: Meanings Score – Delayed Post-Tests

Participant No	Control Group	Soft Experimental Group	Hard Experimental Group
1.	17	22	24
2.	18	19	20
3.	21	20	16
4.	18	20	22
5.	21	25	18
6.	19	19	21
7.	20	22	17
8.	22	21	23
9.	19	17	21
10.	20	21	22
11.	16	21	24
12.	20	20	20
13.	35	21	22
14.	12	20	18
15.	19	23	22
16.	17	18	21
17.	12	16	21
18.	22	23	26
19.	18	21	23
20.	10	18	18
21.	18	20	25
22.	16	22	19

23.	21	19	15
24.	17	23	25
25.	24	18	17
26.	17	23	17
27.	21	7	22
28.	16	24	23
29.	16	17	25
30.	17	20	24
Mean	18.63	20.00	21.03
SD	4.36	3.28	2.98

Running an ANOVA test on the above scores produces the following results:

Comparison of the Three Groups – Delayed Post-Tests

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	-1.367	.928	.309
	Hard	30	-2.400	.928	.030
Soft	PPP	30	1.367	.928	.309
	Hard	30	-1.033	.928	.508
Hard	PPP	30	2.400	.928	.030
	Soft	30	1.033	.928	.508

Continuing the pattern, the P value for the difference between the scores of the soft experimental and the control groups is very high (0.309), indicating that the difference is not significant and therefore, not reliable. The significance value, however, is still very low for the control and hard experimental pair of groups which means that the difference is significant. Finally, the P value for the difference between the hard and soft experimental groups is again very high, meaning that the difference is not significant.

The mean of the scores of the hard experimental group is only 5.02% more than the soft experimental group and 12.1% more than the control group. We can, therefore, say that the participants of the hard experimental group had the highest difference in their retention in comparison with the control group. In addition to this,

the difference in the scores of the soft experimental group and the control group is 7.09% which is not considered significant.

4.3.6.3 Essays.

The general DMs scores in essays by the three groups are given below:

Essays: General DMs Score – Delayed Post-Tests

Participant No	Control Group	Soft Experimental Group	Hard Experimental Group
1.	5	3	10
2.	8	9	9
3.	5	2	8
4.	11	11	3
5.	7	4	3
6.	10	12	10
7.	7	10	9
8.	4	4	9
9.	3	7	18
10.	4	7	15
11.	7	2	10
12.	4	11	6
13.	7	5	19
14.	5	2	8
15.	11	14	7
16.	5	10	13
17.	5	11	13
18.	7	8	14
19.	12	3	3
20.	3	2	4
21.	5	9	4
22.	4	2	6
23.	4	5	3
24.	5	6	6
25.	7	4	2

26.	5	4	9
27.	8	4	11
28.	1	5	17
29.	8	10	13
30.	9	9	10
Mean	6.20	6.50	9.07
SD	2.61	3.57	4.69

For the essay general score, let us run an ANOVA test on the figures:

Comparison of Three Groups – Delayed Post-Tests

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	-.300	.961	.948
	Hard	30	-2.867	.961	.010
Soft	PPP	30	.300	.961	.948
	Hard	30	-2.567	.961	.024
Hard	PPP	30	2.867	.961	.010
	Soft	30	2.567	.961	.024

The pattern continues with the difference between the soft experimental group and the control group being insignificant with the P value of 0.948. However, a break in the pattern in observed, where the P value for the difference between the soft and hard experimental groups is 0.024 which means that there is a significant difference in the scores of the participants. This means that the scores of the participants of the soft experimental group were closer to the control group than they were to the experimental group. The difference between the hard experimental group and the control group continues the pattern of a significant difference with the P value of 0.01. Participants of the hard group scored 33.01% more than the soft experimental group and 37.5% more than the control group. The difference between the control and the soft experimental group was only 4.7%.

For the final statistic, the scores of the three groups for the study-specific DMs in the essays are given below.

Essays: Study DMs Scores – Delayed Post-Tests

Participant No	Control Group	Soft Experimental Group	Hard Experimental Group
1.	4	1	8
2.	6	7	7
3.	3	0	7
4.	4	8	3
5.	4	3	3
6.	9	9	5
7.	5	9	3
8.	2	2	8
9.	1	1	13
10.	3	4	6
11.	7	2	8
12.	3	10	2
13.	3	4	11
14.	2	1	4
15.	8	8	4
16.	5	2	10
17.	3	8	9
18.	5	5	10
19.	7	3	1
20.	3	2	3
21.	4	8	3
22.	4	1	5
23.	4	3	2
24.	4	3	4
25.	5	4	2
26.	1	4	7
27.	2	4	8
28.	1	3	13
29.	7	2	10
30.	7	9	9

Mean	4.20	4.33	6.27
SD	2.10	2.98	3.41

An ANOVA test in SPSS provides the following results:

Comparison of the Three Groups – Delayed Post-Tests

Groups	Comparison with	N	Mean Difference	Std. Error	Sig.
PPP	Soft	30	0.133	.746	.983
	Hard	30	2.067	.746	.019
Soft	PPP	30	0.133	.746	.983
	Hard	30	1.933	.746	.030
Hard	PPP	30	2.067	.746	.019
	Soft	30	1.933	.746	.030

Study-specific DMs scores reflect the trends in the general DMs score. The statistical test again indicates a very high P value of 0.983 for the difference between the control and the soft experimental group, which implies that the difference is not significant. The percentage difference is only 3.04% with the soft experimental group having the edge; however, this difference cannot be considered significant due to the high P value. The significance value for the difference between the scores of the hard experimental and the other two groups is, however, low (0.19 for the control group and 0.30 for the experimental group), indicating a significant difference. The percentage difference comes out to participants of the hard experimental group having scored 39.54% and 36.6% more than the participants of the control group and the soft experimental group, respectively.

Examining the total share of the study-specific DMs from the total number of DMs in essays in the post and delayed post-tests, it is revealed that in the post-test, the study-specific DMs made up 64.8% of the total DM use by the participants of the control group, while this percentage was 68.63% for the soft experimental group. In the delayed post-test, the percentage for the control group actually increased to 67.7% and decreased to 66.66% for the soft experimental group. For the hard experimental group, this ratio was 68.57% in the post-test and remained almost consistent with the percentage being 69.11%.

4.3.6.4 Discussion.

The delayed post-tests were by far the most interesting part of the current study where the previous trends in the scores of the groups broke and long term retention of the concepts was tested. The results revealed that the soft DDL method and the PPP method were similar in their long term benefits for the learners, because the difference between their scores became statistically insignificant. The delayed post-test results almost became equal to the level of the pre-tests and were found statistically similar upon exploration. Hard DDL, however, showed sustained improvement in the results even after the 3-month delayed post-test. There are no publicly available studies, which compare the effectiveness of Hard DDL and Soft DDL in delayed post-test, including Chujo et al. (2012), which only administered a post-test.

The higher scores of the Hard DDL over the longer period can be explained by its edge in involvement load, noticing, input enhancement and stimulation of cognitive abilities, all of which are considered to be useful for language learning (Hulstijn & Laufer, 2001; R. W. Schmidt, 1990; Smith, 1993). This is supported by Yanagisawa and Webb's (2022) meta-analysis of studies showing that higher involvement load usually leads to better outcomes for the learners. Based on the factors that differentiate hard DDL from the other two methods, it can be said that it is more effective in providing long term benefits for learners in learning how to use discourse markers. These findings are in line with the constructivist thought which believes higher levels of engagement and the resulting problem solving and meaning making processes to be the basis for language learning.

4.4 Summary

The chapter explored the status of the three groups at various stages before and after the treatment. It was found that the groups were similar before the receiving the treatment from the study, providing an ideal base for providing treatment and conducting research. The Hard DDL and Soft DDL groups showed markedly higher improvement over the control group in the post-tests, which can be credited to their higher involvement load, input enhancement, higher levels of noticing and stimulation of cognitive skills. The improvements, however, did not remain the same in the delayed post-test conducted after 3 months, where the Hard DDL still had a

considerable edge over the control group but the Soft DDL group had become almost equal with the control group in comparison with the pre-test scores. This indicated that the Hard DDL method provides better results immediately as well as in long term, making it more effective in the teaching of DDL.

CHAPTER 5

DISCUSSION AND CONCLUSION

While the data obtained in the study has already been analyzed in the previous research along with a brief discussion, it is now time to interpret the results and connect them with the general body of knowledge related to the components of this study. The purpose of this chapter is to summarize the findings, discuss them in light of previous work on the topic, provide suggestions and recommendations, and conclude this study.

This study aimed to measure the effectiveness of the two versions of the Data-Driven Learning Method within the context of Pakistan. It compared the results of the application of these methods with a method that is traditionally used in Pakistani classrooms, i.e. the Presentation, Practice, Production (PPP) model. It did so by employing two different instruments: the sentence-making exercise and the essay-writing exercise. The purpose of the essay writing exercise was to see if the treatments given to them had any effect on how they used the language. The study examined if there was only improvement in the DMs taught in the study or if it had a larger, wider effect too. Similarly, the study also used a sentence-making exercise. The reason such an exercise was used was to directly gauge any change in understanding of the DMs and how the learners use them. This involved not only assigning them a score out of a set total score but also seeing how differently they used each DM. Finally, the study aimed to find out how the learners retained the concepts over a long time, for which a delayed post-test was conducted.

5.1 Summary of the Findings

To increase the credibility of the results, this study used triangulation to collect the data. There were two instruments that served to collect data: the sentence-making exercise as well as the essays. The cumulative scores of the two groups show us that the two experimental groups were able to provide better overall scores than the control group in the post-test, where the control group showed a 13.61% improvement, while the HEG showed a 28.99% improvement and the SEG showed

25.16% improvement. The difference between the HEG and the SEG is considered to be statistically insignificant, therefore giving us reason to believe that they can show similar outcomes for learners.

A slightly different picture shows up upon the examination of the cumulative scores in the delayed post-tests, which shows a decrease in the scores of all three groups. The decline in scores was mostly consistent with the control group seeing slightly more decline than the other two in cumulative scores. However, since the HEG had higher scores in the post-test, the overall improvement in comparison with the pre-test remained the highest with 13.80% higher scores. In comparison, the control group actually showed a little bit of a decline in the score by 3.1%, which can very well be a chance occurrence. The SEG showed a 4.05% improvement in comparison, therefore being second to HEG for long-term retention of the concepts. This provides evidence to believe that the HEG is overall the better methodology for instilment of concepts for long-term usage in learners.

5.1.1 Findings from the Sentence Making Exercise

The sentence-making exercise looked at the use of DMs in two ways. The first was to see if the learners were able to make three different sentences correctly with the given DMs. The second part of it was to see how many different ways they had used each of the DMs. It was found that the participants of the HEG showed higher levels of improvement than the control group, when it came to the ability to form sentences in general without any focus on their meanings. The mean of scores in the post-test was 9.49% more than the control group, while the mean of scores for SEG was 7.93% more than the control group. This demonstrated to us that both the HEG and SEG were able to provide better scores than the control group. However, the results posted by SEG carry a low P value, which makes it likely that the results were by chance. The difference in improvement of the scores of the participants of the HEG and the SEG in comparison with the control group becomes more pronounced upon looking at the different ways that the DMs were used. This allows the inference that the two experimental groups were more effective in teaching DMs in comparison with the treatment of the control group. Furthermore, it can also be seen that the HEG has a consistent edge in both scores in comparison with the SEG, leading to the belief

that it is the more effective of the two. However, solely looking at the figures, it is observed that the difference between the two groups was not statistically significant.

If the long-term retention of the concepts is considered, it is observed that the trend in the post-test only partially continues to reflect in the delayed post-test. While the HEG does also show a decline in scores like the other two groups, it still posted an 11.29% better result than the pre-tests. SEG in contrast only retained 8.78% better scores, however, the results are not considered statistically significant. For the score indicating the diversity in using the DMs, HEG had even more contrast with the rest of the groups with 12.1% better scores. Meanwhile, SEG showed only 7.09% higher scores, however, the P value for SEG is too high for it to be considered a significant result. These results indicate that for the overall retention, HEG was able to provide more improvement in scores for the learners.

5.1.2 Findings from the Essays

The analysis of the essays looked at two different things: the number of DMs used overall and the number of DMs taught in the study. The results of the study found that the participants of both of the experimental groups showed higher percentage scores than the control group in the number of DMs used in their essays. Participants of the HEG used 29.1% more DMs than the control group in general. In comparison, this figure was 27.48% for the SEG. However, while this does present as a high percentage difference, due to the nature of statistics, this difference is not considered very significant. The real difference comes when you look at the number of the study DMs used by each group where we find that the participants of the HEG scored 34.61% and the participants of SEG scored 32.62% more than the control group, both of which are significant figures. This reaffirms two of our previous findings: the first is that both of the experimental groups showed higher levels of improvement than the control group. The second finding is that the participants of the HEG have only slightly higher scores than the SEG, which is not enough to justify a significant difference.

Coming to long-term retention, the trends appear to be similar to the sentence-making exercise. All three groups showed a decline in scores, however, the scores of the HEG were still 37.5% higher for general DMs and 39.54% better for study-

specific DMs. The SEG only showed 4.7% higher scores in general DMs and a 3.04% higher score in study-specific DMs, both of which are statistically insignificant. This implies that the learners learning through the PPP method and the SEG retained almost the same number of concepts over the long term. To sum up, HEG had better long-term outcomes for learners than the other two groups, both of which showed similar levels in the delayed post-test.

5.2 Discussion

The current study employed triangulation to make the results more credible. The current study found similar results from both research instruments. The sentence-making exercise directly tested the knowledge of the use of DMs while essays tested the actual application of the DMs in real-world usage of the language. Sentence making exercise also provided extra data that was not influenced by the topic or the length of the essays written by the students, making the results more reliable.

Overall, versions of Data-Driven Learning were able to provide better outcomes for learners, similar to studies like Corino and Onesti (2019), Boontam and Phoocharoensil (2018), Yilmaz and Soruc (2015) and Barabadi and Khajavi (2017), all of which experimented with teaching some sort of vocabulary item. We saw some consistent trends in the results of the three groups; however, there were also some interesting exceptions. HEG showed significantly higher results overall in the post-tests as well as the delayed post-test. While the results of the SEG were closer to the HEG and showed significant improvement over the control group in the post-tests, this changed in the delayed post-tests when the scores of SEG almost equalized with the control group. These findings further substantiate the theoretical framework of the current study, which provides many bases for DDL activities to be considered effective. As has been discussed in length about DDL previously, the inherent nature of DDL is believed to facilitate learners in noticing language patterns as well as having higher involvement load, which directly contribute to having higher learner outcomes (Schmidt, 2012). The questions asked during the DDL activities also play into several of these concepts by raising consciousness and engaging the problem solving and cognitive skills of the learners, as has been discussed by Flowerdew (2015) and Meunier (2020).

The increase in the post-test results of learners is highly reflective of the findings of the highly influential meta-analysis of research on DDL by Boulton and Cobb (2017, cited in Crosthwaite, 2019). Overall, DDL also managed to post higher levels of improvement in the delayed post-test, when it seemed that the control group did not retain anything at all. The decline in the delayed post-test scores can also be seen in the study conducted by Yao (2019), however, similar to this study, it can be seen that since Hard DDL resulted in higher scores in the post-test, the delayed post-test scores remained high too despite the decline. This substantiates the findings of this study where the hard version of DDL was able to provide sustained higher results even after three months, when the other two groups showed decline to a level, where there was no significant difference from the pre-test.

It was found that the hard version of DDL provided higher improvement in both of the tests employed in the study all across the different components. The results from the administration of the soft version of DDL were more varied in comparison. There is very little empirical research about the difference between the results of hard Data-Driven Learning and Soft Data-Driven Learning. The available research, however, is a bit divisive with some considering both to be equally effective while others consider hard DDL to be slightly more effective, e.g. Chujo et al (2012) and Uchibori (2013). While it showed significant improvement in the learners' ability to form sentences using the DMs, it did not show much improvement in terms of the different ways they used them. The results from the soft version of DDL seemed to lie between those of the control group and the HEG, while the control group constantly scored under the other two. The participants of the SEG showed higher scores than the control group in every metric in the post-test, with the only exception being the general ability to use the DMs in sentences, where the increase in scores had a lower significance value. This, however, still demonstrates a greater improvement in scores of the learners than the control group, as we observe in the results of the analysis of the essay, where they show comparable improvement to the HEG.

In trying to find a rationale for these results, the theoretical framework of the research provides a source to find answers to the reasons behind findings of this study. The suggestions of Schmidt (1990) about the facilitative role of noticing in language lead us to believe that the component of drawing attention to language

patterns in both versions of the DDL leads to higher results. This is echoed in the study conducted by Corino and Onesti (2019) who report that learners administered with DDL demonstrated increased motivation and involvement and consequently higher results in learning. Learners administered with DDL also enjoy higher levels of cognitive stimulation. Learners had to analyze the instances of the use of DMs, guess their usage from the context, predict how they will be used in more contexts and then produce new sentences. The presence of different types of use of the same DMs also meant that learners had to compare their meanings, hypothesize their differences and then draw conclusions, which were then reflected in the form of their usage. These processes involved in learning by DDL methods have been theorized by O'Sullivan (2007) to increase language learning. Barabadi and Khajavi (2017) also attribute the higher scores of learners using DDL to their active role in their own learning, which involves discovering the rules on their own and using their inductive skills.

The higher levels of improvement by the participants of the HEG can be attributed to certain features that increase noticing, enhance the input and stimulate the cognitive skills more than SEG. The procedure of the hard DDL method involves the interaction of learners with computers. This provides two advantages over soft DDL. The first is that the learners have access to an even larger number of instances where the DM is used, providing more contexts and examples for learners to analyze. The second advantage is that learners can actively interact with the software; they can click on the instances of the use of the DM and find even more context. This is also reflected in the findings presented by Abuczki, Parmaxi and Nicolaou (2018) who report that learners find the use of digital tools more enjoyable, thus showing higher improvement in learning DMs. The features of a computer-enabled DDL method mean that learners not only get more input, but also feel more interested in the content. This means that they feel the need to interact more with the software and then they have to actively find and evaluate the instances, which are the essentials of the Involvement Load Hypothesis by Hulstijn and Laufer (2001).

Another important thing to note is that learners showed higher levels of difference in improvement in scores in the actual application of the DMs in essays than in just making the sentences. This means that while the learners being treated with the PPP method might be able to learn the concepts, they have a harder time

incorporating them into their actual language use. The HEG and the SEG showed that the participants who studied using the two versions of DDL were able to improve their actual real-life practical usage of the concepts.

Coming to the delayed post-tests, they demonstrated an increased gulf between the HEG and the SEG, where the participants of the HEG were able to retain their ability to use the DMs even after three months, while in comparison the SEG and the control group were almost at the level of the pre-tests. The decline in results was also seen by Yao (2019) who noticed an almost 35% decrease in the means. These findings are consistent with the findings of the current study, where the delayed post-test scores dropped significantly. There is, however, little empirical work on the comparison of the two versions of DDL.

It has been observed earlier in the current study that the results were not significantly different between the HEG and the SEG in the post-test, which brings up the question of why there is such a significant difference between the two groups. The answer seems to lie in the way they are actually presented to the learners, i.e. through a computer versus through paper handouts. It can be concluded from these results that the higher the involvement load as well as cognitive stimulation is, the higher long-term retention there will be in learners as well as the ability to apply the learned concepts in actual usage of language. While the soft DDL method did have interactive and student-driven learning, it was severely eclipsed by the hard DDL in these components. Paper handouts might be considered as plain, boring and a chore by the learners, while hard DDL excites interest as well as provides more opportunities for discovery by venturing into the context as well as looking at instances, since it is not limited by the paper space. These postulations are in line with our earlier observations about involvement load and the stimulation of cognitive skills.

The results of the study provide grounds for the belief that both versions of DDL are not only suitable for learners at higher secondary level in Pakistan, but that, learners can be provided autonomy in their learning, since DDL is a learning method that heavily relies on learners deriving out concepts. Pakistani language learners are able to engage in inferential learning, as is the case with Data Driven Learning, and can produce better results than traditional forms of learning. This is in line with the

findings of other researchers, albeit in different settings, that learners are able to take charge of their learning and can form concepts on their own (e.g. Nazari, 2014, Wachob, 2005)

5.3 Conclusion

The conclusions of this research are based on quantitative results derived from the data obtained during the study using the two research instruments. These instruments were administered to the participants of the study before any treatment, which revealed that the three groups of participants were at equal levels. Following the pre-tests, the three groups were administered three different teaching methods: the PPP method to the control group and the soft and hard DDL methods to the two experimental groups. These groups were then tested following treatment in an immediate post-test and then again in a delayed post-test after three months had passed by. The results of these three tests were, then, used to test the hypotheses put forth by this study.

Hypothesis 1 dealt with the comparison of the outcomes of the hard DDL method with the PPP method. It was found that the learners using the hard DDL method outperformed the participants of the PPP group in every single component of the tests and scored significantly higher.

Hypothesis 2 was related to the comparison of the SEG with the control group. We again found similar results to the HEG, in which the participants learning through soft DDL scored higher than their counterparts using the PPP method. The results were mostly consistent except for the insignificant improvement in the general ability to form sentences with the DMs.

The purpose of hypothesis 3 was to compare the post-test results of soft and hard versions of the DDL method. The results show that the post-test results of the two versions of the method are indeed similar. While we do see a higher consistent percentage of results in the participants learning through hard DDL, the results are statistically insignificant ($P > 0.05$). Therefore, it can be concluded that the two groups give similar results in the post-test, albeit hard DDL might have a slight edge over soft DDL.

Finally, hypothesis 4 deals with the long-term retention of concepts by the participants of all three groups. The findings showed us that the long-term results of the three groups are not similar with the hard DDL showing significantly more retention than the other two groups. Participants learning through Soft DDL unexpectedly dwindled in scores and were only slightly above the participants of the SEG, which leads us to believe that higher levels of involvement load as well as motivation and interest in hard DDL lead to better long-term retention in learners.

5.4 Recommendations

The current study culminates in recommendations for teachers, education planners and language learners, which are perhaps the most important part of the conclusion of the study, as they provide practical implications of the study.

5.4.1 Recommendations for Teachers

The following recommendations can be made for teachers in view of the above results:

1. Teachers should consider implementing versions of Data Driven Learning in classrooms. The results of the current study clearly make a case for the Data Driven Learning approach, although more so for the hard version of the approach, compared to the soft paper-based version. In view of this, the hard version of the approach may be implemented wherever possible; otherwise, the soft version may still provide an edge over the conventional PPP method. It can be recommended that DDL may be made a part of regular classroom activities to target linguistic items, such as DMs, which present an unusual nature. Application of the method for other linguistic items may be subject to the results of the relevant research, however, the results of the current study do suggest a benefit to its application for teaching DMs.
2. Teachers should consider prioritizing durable learning in class. The current study showed that while the results shown by the three groups may have been close in some categories in the post-test, the difference increased in the delayed post-test. This implies that while there may be immediate results of a teaching methodology, they might not be sustained for longer periods of time, rendering the teaching essentially ineffective. DDL, in this regard, showed an edge over the traditional PPP method and the learners treated with the hard

DDL method showed better long-term retention, therefore, durable learning methods, such as the DDL, may be considered for application in classrooms.

3. Teachers should consider balancing different teaching methods in classrooms. While the hard version of DDL produced better results overall, the soft version of DDL still proved its merit in some measures of performance, especially in the post-test. A combination of both, so in order to not to strain the available resources or in the case of their lack of their wider availability, may be employed so that the overall results may improve.

5.4.2 Recommendations for Educational Planners

While the primary participants of the educational process are teachers and learners, educational planner play a huge role in how education takes shape in classrooms. In view of this and the above study, the following recommendations can be made for educational planners.

1. Data Driven Learning approaches can be considered by educational planners to be adopted into regular education. As far as the scope of this study is concerned, DDL may be especially beneficial to irregular linguistic items such as DMs. Incorporating DDL into regular classrooms allows teachers to fill the gaps in learners' understanding and ensure durable learning. In addition, since both versions of the DDL had varying levels of success, the two versions can be blended together to provide activities that are feasible as well as interesting.
2. Language teachers could benefit from standardized training in DDL. The DDL method significantly differs from the traditional PPP method. While some teachers may be able to explore and adopt the practices into their classrooms of their own volition, training for the method must be standardized in order for teachers to consistently provide learners with DDL activities in order to boost their learning.
3. Educational planners may seek to include more technology into language learning. The current study demonstrated higher results for both versions of DDL, however, the technologically richer version, the hard DDL, stood out as more effective. This implies that technologically enhanced classrooms may be more effective at providing higher outcomes for language learners.
4. The current study has shown that there is merit in allowing learners to be independent and work as critical thinkers. This should be reflected in learning

curriculums, by including activities that enable learners to explore language rules and structures on their own as well as that promote critical thinking.

5. Long term retention should be kept in mind when evaluating any teaching methodology, as it shows that the process of education has been successful in helping learners to internalize concepts and make them a part of their permanent knowledge base.

5.4.3 Recommendations for Students

In view of the above study, the following recommendations can be made for students:

1. Students should try to take charge of their own learning and should be confident in inferring rules. However, they must also consult their teachers and take and implement any feedback. The process of learning should be treated as a process of trial and error, where learning only takes place by attempting to understand something first.
2. Students should actively engage with DDL activities and materials. DDL can look intimidating from the outside; however, it has the potential to be highly engaging for the learners. Students, therefore, should approach it with openness and interest, no matter if the DDL activities are paper based or computer based. This will allow them to find out patterns and rules within language on their own and create deeper understanding.
3. The goals of language learning go beyond any exam or test, therefore, long term learning should be the main focus of any learning, as language is a tool to be used all lifelong. The actual use of language mostly happens outside a classroom where it is used to communicate with the people around us. Students should, therefore, seek to adopt DDL activities in their learning, as it can allow them to develop concepts and understand language rules as they are used in real life that will stay with them for beyond the immediate results, empowering the to use language anytime and in any situation.
4. Understanding language use should be preferred over rote memorization of rules. Rote memorization may have some immediate results, however, the learning in not durable and does not reflect in enhanced real-life use. Hence, learners should try to understand and infer linguistic patterns and rules, in order to get a grip on the actual use of language.

5.5 Limitations of the Study

As any research does, the current study also had a few limitations. The very first thing to note here is that the research was conducted in a higher-income school with learners from mostly middle class or upper middle class. Learners from higher socio-economic backgrounds have been found to have an advantage over learners from poor socio-economic backgrounds (Akram et al., 2021). Learners studying in high income private schools in Pakistan also tend to do better in language learning, thanks to higher level of resources available to them as well as acquisition-rich backgrounds at home (Shamim, 2017). Hence, these learners are generally higher in their proficiency levels than learners from schools with lower income backgrounds. Another limitation was the time available to the researcher to implement the treatment due to restrictions in the coursework time of the learners. Educational programs at higher secondary levels are very rigorous due to the exams conducted by Board of Intermediate and Secondary Education, following a very tight schedule in order to cover course and compete. Therefore, schools were not willing to allow time for a long study to be conducted.

Finally, the study was also limited by the number of participants in the study. While consent was obtained from a large number of learners, only a part of them returned all of the required test samples. More schools could not be contacted due to lack of feasibility as well as the schedule of these learners, as they had to prepare for their board exams. Therefore, the researcher had to settle for fewer samples than ideal.

5.6 Future Research

The current study revealed many interesting things that may merit further inquiry. The following recommendations can be made regarding further avenues of research:

1. There was a consistent and obvious, albeit small, difference between the results of soft DDL method and the hard DDL method. Further research into the reason for this difference can be more revealing.

2. Research could look into why there was a significant decline in the results of the soft DDL method in comparison with the hard DDL method in the delayed post-tests.
3. The test mixed both boys and girls. It may be tested if there are any gender differences which may lead to different results in the study methods.

REFERENCES

Abuczki, A., Parmaxi, A., & Nicolaou, A. (2018). *Teaching Discourse Markers in a Technologically-Enhanced Language Classroom* (pp. 323–336). https://doi.org/10.1007/978-3-319-91743-6_25

Aidinlou, N. A., & Mehr, H. S. (2012). The Effect of Discourse Markers Instruction on EFL Learners' Writing. *World Journal of Education*, *2*(2), 10–16.

Aijmer, K. (1996). *Conversational Routines in English: Convention and Creativity*. Longman.

Aijmer, K. (2002). English Discourse Particles. *English Discourse Particles*, 1–315.

Akram, D. M., Ahmad, S., Ahmad, S., Javed, M., Divaya, A., & Naeem, D. (2021). Role of Socioeconomic Status in ESL Learning. *Turkish Online Journal of Qualitative Inquiry*, *12*(8), Article 8.

Alraddadi, B. M. J. (2016). The effect of teaching structural discourse markers in an EFL classroom setting. *English Language Teaching*, *9*(7).

Also. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/also

ALSO definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/also

Asl, H. D., & Moradinejad, A. (2016). The Effect of Explicit Instruction of Discourse Markers on Iranian EFL Learners' Speaking Ability. *Journal of Applied Linguistics and Language Research*, *3*(5), Article 5.

Bao, K. (2021). *Should We Use It in Our Classrooms: An Analysis of Data-Driven Learning Research*. https://www.semanticscholar.org/paper/Should-We-Use-

It-in-Our-Classrooms%3A-An-Analysis-of

Bao/c16654e86b8ccf1e8bb915e043409f108d7adad3

Barabadi, E. 1, & Khajavi, Y. 2 1 T. (2017). *The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary*. https://doi.org/10.1080/2331186X.2017.1283876

Because. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/because

BECAUSE definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/because

Bennett, G. (2010). *Using Corpora in the Language Learning Classroom*. https://www.academia.edu/70735084/Using_Corpora_in_the_Language_Learning_Classroom

Blakemore, D. (1987). *Semantic Constraints on Relevance*. Blackwell.

Blakemore, D. (1992). *Understanding utterances*. Blackwell.

Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English Preposition Learning through Data-Driven Learning (DDL). *3L The Southeast Asian Journal of English Language Studies*, *24*, 125–141. https://doi.org/10.17576/3L-2018-2403-10

Boulton, A. (2007). DDL is in the details... And in the big themes. *Corpus Linguistics Conference: CL2007*, XX. https://hal.science/hal-00326994

Boulton, A. (2009). Data-driven learning: On paper, in practice. *Corpus Linguistics in Language Teaching.*, 17–52.

Boulton, A. (2010). Data-Driven Learning: Taking the Computer Out of the Equation. *Language Learning*, *60*. https://doi.org/10.1111/j.1467-9922.2010.00566.x

Boulton, A. (2011). Data-driven learning: The perpetual enigma. *Explorations across Languages and Corpora*.

Boulton, A., & Landure, C. (2016). Using Corpora in Language Teaching, Learning and Use. *Recherche et Pratiques Pédagogiques En Langues de Spécialité. Cahiers de l'Apliut*, *Vol. 35 N° 2*, Article Vol. 35 N° 2. https://doi.org/10.4000/apliut.5433

BUT definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/but

Chang, P. (2012). Using a stance corpus to learn about effective authorial stance-taking: A textlinguistic approach. *ReCALL*, *24*(2), 209–236. https://doi.org/10.1017/S0958344012000079

Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-Based, Computer-Based, and Combined Data-Driven Learning Using a Web-Based Concordancer. *Language Education in Asia*, *3*, 132–145. https://doi.org/10.5746/LEiA/12/V3/I2/A02/Chujo_Anthony_Oghigian_Uchibori

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, *25*(3), 301–315.

Corino, E., & Onesti, C. (2019). Data-Driven Learning: A Scaffolding Methodology for CLIL and LSP Teaching and Learning. *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00007

Corpus Linguistics for ELT: Research and Practice—1st Edition—Ivor. (n.d.). Retrieved November 21, 2023, from https://www.routledge.com/Corpus-Linguistics-for-ELT-Research-and-Practice/Timmis/p/book/9780415747127

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson.

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Crosthwaite, P. (Ed.). (2019). Data-driven learning and younger learners: Introduction to the volume. In *Data-Driven Learning for the Next Generation*. Routledge.

Crystal, D. (2011). *A Dictionary of Linguistics and Phonetics*. John Wiley & Sons.

Definition of HOWEVER. (n.d.). Retrieved April 22, 2024, from https://www.merriam-webster.com/dictionary/however

Definition of NOW. (2024, April 22). https://www.merriam-webster.com/dictionary/now

Definition of THOUGH. (n.d.). Retrieved April 22, 2024, from https://www.merriam-webster.com/dictionary/though

Ellis, R. (1993). The Structural Syllabus and Second Language Acquisition. *TESOL Quarterly*, *27*(1), 91–113. https://doi.org/10.2307/3586953

Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning*, *1*(1), 80–109. https://doi.org/10.1515/jccall-2021-2004

Flowerdew, L. (2015). Whither the twain shall meet: Data-driven learning and language learning theories. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple Affordances of Language Corpora for Data-driven Learning* (pp. 15–36). John Benjamins Publishing Company. https://doi.org/10.1075/scl.69.02flo

Fraser, B. (1993). *Discourse Markers across Language*. https://eric.ed.gov/?id=ED396548

Fraser, B. (1996a). Pragmatic markers. *Pragmatics*, *6*(2), 167–190. https://doi.org/10.1075/prag.6.2.03fra

Fraser, B. (1996b). Pragmatic Markers. *Pragmatics; Vol 6, No 2 (1996)*, *6*. https://doi.org/10.1075/prag.6.2.03fra

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, *31*(7), 931–952. https://doi.org/10.1016/S0378-2166(98)00101-5

Fung, L., & Carter, R. (2007). Discourse Markers and Spoken English: Native and Learner Use in Pedagogic Settings. *Applied Linguistics*, *28*(3), 410–439. https://doi.org/10.1093/applin/amm030

Fung, L. P. (2003). *The use and teaching of discourse markers in Hong Kong: Students' production and teachers' perspectives* [Thesis (University of Nottingham only)]. University of Nottingham. https://eprints.nottingham.ac.uk/11437/

Gabrielatos, C. (2005). Corpora and Language Teaching: Just a Fling or Wedding Bells? *TESL-EJ*, *8*(4). https://eric.ed.gov/?id=EJ1068106

Gilquin, G., & Granger, S. (2010). *How can data-driven learning be used in language teaching?* (pp. 359–370).

Granger, G. G., Sylviane. (2010). How can data-driven learning be used in language teaching? In *The Routledge Handbook of Corpus Linguistics*. Routledge.

Haberlandt, K. (1982). Reader Expectations in Text Comprehension. In J.-F. Le Ny & W. Kintsch (Eds.), *Advances in Psychology* (Vol. 9, pp. 239–249). North-Holland. https://doi.org/10.1016/S0166-4115(09)60055-8

Hadley, G. (2002). An Introduction To Data-Driven Learning. *RELC Journal*, *33*(2), 99–124. https://doi.org/10.1177/003368820203300205

Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, *15*(2), 159–182. https://doi.org/10.1177/1362168810388694

However. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/however

HOWEVER definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/however

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, *26*(2), 163–183. https://doi.org/10.1017/S0958344014000020

Hulstijn, J. H., & Laufer, B. (2001). Some Empirical Evidence for the Involvement Load Hypothesis in Vocabulary Acquisition. *Language Learning*, *51*(3), 539–558. https://doi.org/10.1111/0023-8333.00164

I THINK definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/i-think

Instead. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/instead

Jabeen, F., Rai, A., & Arif, S. (2011). A corpus based study of discourse markers in British and Pakistani speech. *International Journal of Language Studies*, *5*, 69–86.

Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, *14*(2), 151–162. https://doi.org/10.1016/0346-251X(86)90004-7

Johns, T. (1991). *SHOULD YOU BE PERSUADED - TWO SAMPLES OF DATA-DRIVEN LEARNING MATERIALS*.

Johns, T. (1993). Data-driven learning: An update. *ᵀᴱᴸL & CALL*, *3*, 23–32.

Johns, T. (1998). Contexts: The Background, Development and Trialling of a Concordance-based CALL Program. In *Teaching and Language Corpora*. Routledge.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In *Teaching and learning by doing corpus analysis* (pp. 105–117). Brill.

Johns, T. F. (1991). From Printout to Handout: Grammar and Vocabulary Teaching in the context of Data-driven Learning in Johns, TF and P. *King (Eds.)*.

Jones, C. (n.d.). *Spoken discourse markers and English language teaching: Practices and*.

Jones, C. (2011, December 15). *Spoken discourse markers and English language teaching: Practices and pedagogies* [Thesis (University of Nottingham only)]. University of Nottingham. https://eprints.nottingham.ac.uk/12260/

Jones, C., & Carter, R. (2014). Teaching spoken discourse markers explicitly: A comparison of III and PPP. *International Journal of English Studies*, *14*(1), Article 1. https://doi.org/10.6018/ijes/14/1/161001

Jucker, A. H. |Ziv. (1998). Discourse Markers. In *Pbns.57*. John Benjamins Publishing Company. https://benjamins.com/catalog/pbns.57

KAMALI, F., & NOORI, H. (2015). The impact of discourse markers instruction on improving writing of intermediate EFL learners. *Cumhuriyet Üniversitesi Fen Edebiyat Fakültesi Fen Bilimleri Dergisi*, *36*(3), 944–949.

Kapranov, O. (2018). *The Impact of Implicit Instruction upon the Use of English Discourse Markers in Written Tasks at the Advanced Beginners' Level of EFL Proficiency.*

Kennedy, G. (2014). *An Introduction to Corpus Linguistics*. Routledge.

Khandaghi Khameneh, A., & Fakhraee Faruji, L. (2020). The Effect of teaching Discourse Markers (DMs) on speaking achievement among Iranian intermediate EFL learners. *International Journal of Research in English Education*, *5*(4), 1–13.

Klomkaew, T., & Boontam, P. (2023). Effectiveness of Grammar Learning through Data-Driven Learning (DDL) with Undergraduate Students. *Shanlax International Journal of Education*, *11*(4), 30–44.

Lew, R., & Jackson, H. (2013). *The Bloomsbury Companion to Lexicography*.

Lin, M. (2021). Effects of Data-Driven Learning on College Students of Different Grammar Proficiencies: A Preliminary Empirical Assessment in EFL Classes. *SAGE Open*, *11*, 215824402110299. https://doi.org/10.1177/21582440211029936

Lüdeling, A., & Kytö, M. (2009). Corpus linguistics: An international handbook. In *Corpus Linguistics: An International Handbook* (p. 1353).

Malik, S., Waheed, D. S., & Bhatti, D. Z. I. (2021). A Corpus Based Analysis of Discourse Markers in Pakistani and American English: New Digital Mass Media Era. *International Journal of Innovation*, *15*(4).

Maschler, Y., & Schiffrin, D. (2015). Discourse Markers Language, Meaning, and Context. In *The Handbook of Discourse Analysis* (pp. 189–221). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118584194.ch9

McEnery, T., & Wilson, A. (1997). Teaching and Language Corpora(TALC). *ReCALL*, *9*(1), 5–14. https://doi.org/10.1017/S0958344000004572

McEnery, T., & Wilson, A. (2001). *McEnery, T. & Wilson, A. (2001). Corpus Linguistics (second edition)*.

McKay, S. (1980). Teaching the Syntactic, Semantic and Pragmatic Dimensions of Verbs. *TESOL Quarterly*, *14*(1), 17–26. https://doi.org/10.2307/3586805

Meunier, F. (2020). Data-Driven Learning: From Classroom Scaffolding to Sustainable Practices. *EL.LE*, *2*, JournalArticle_2968. https://doi.org/10.30687/ELLE/2280-6792/2019/02/010

Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction*. Cambridge University Press. https://doi.org/10.1017/CBO9780511606311

Moghaddasi, M. E., Bavali, M., & Behjat, F. (2020). How Does Explicit and Implicit Instruction of Formal Meta-discourse Markers Affect Learners' Oral Proficiency? *International Journal of Foreign Language Teaching and Research*, *8*(33), 93–108.

Nawaz, D. S., Rehman, W., Tanoli, Z. A., & Abdullah, M. (2021). The Use of Discourse Markers by Pakistani ESL Learners in Their Explanatory and Reflective Essays. *Jahan-e-Tahqeeq*, *4*(4).

Noor, M. (2021). Nativisation And Variation Of Discourse Markers In Pakistani English. *University of Chitral Journal of Linguistics and Literature*, *5*(II), Article II. https://doi.org/10.33195/46xt8c56

NOW definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/now

Nurmukhamedov, U. (2012). Online English-English Learner Dictionaries Boost Word Learning. *English Teaching Forum*, *50*(4), 10–15.

Oh. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/oh

OH definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/oh

Östman, J.-O. (1981). *"You Know": A discourse-functional study*. John Benjamins Publishing.

O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, *19*(3), 269–286. https://doi.org/10.1017/S095834400700033X

Papaioannou, V. (2018). *Teaching english as a foreign language through a data—Driven learning perspective using an annotated pedagogic corpus of english text books in a greek high school class*

PERHAPS definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/perhaps

Rahimi, F., & Riasati, M. J. (2012). The Effect of Explicit Instruction of Discourse Markers on the Quality of Oral Output. *International Journal of Applied Linguistics and English Literature*, *1*(1), Article 1. https://doi.org/10.7575/ijalel.v.1n.1p.70

Rao, V. C. S. (2018). The importance of collocations in teaching of vocabulary. *Journal of research scholars and professionals of English language teaching*, *7*(2), 1-8.

Really. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/really

REALLY definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/really

Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics*, *29*(6), 1139–1172. https://doi.org/10.1515/ling.1991.29.6.1139

Roberts, P., & Priest, H. (2006). Reliability and validity in research. *Nursing Standard*, *20*(44), 41–46.

Robinson, P. (1995). Attention, Memory, and the "Noticing" Hypothesis. *Language Learning*, *45*(2), 283–331. https://doi.org/10.1111/j.1467-1770.1995.tb00441.x

Rutherford, W. E. (1987). The meaning of grammatical consciousness-raising. *World Englishes*, *6*(3), 209–216. https://doi.org/10.1111/j.1467-971X.1987.tb00201.x

Rutherford, W. E., & Smith, M. S. (1985). Consciousness-raising and Universal Grammar. *Applied Linguistics*, *6*(3), 274–282. https://doi.org/10.1093/applin/6.3.274

Sadeghi, B., & Kargar, A. (2014). The effect of explicit instruction of discourse markers on EFL learners' writing ability. *International Journal of Educational Investigations*, *1*(1), 328–338.

Schiffrin, D. (1987). *Discourse Markers*. Cambridge University Press.

Schmidt, R. (2012). Chapter 2. Attention, awareness, and individual differences in language learning. In W. M. Chan, K. N. Chin, S. Bhatt, & I. Walker (Eds.), *Perspectives on Individual Characteristics and Foreign Language Education* (pp. 27–50). De Gruyter Mouton. https://doi.org/10.1515/9781614510932.27

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning1. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.

Schourup, L. (1999). Discourse markers. *Lingua*, *107*(3), 227–265. https://doi.org/10.1016/S0024-3841(96)90026-1

Shaffer, C. (1989). A Comparison of Inductive and Deductive Approaches to Teaching Foreign Languages. *The Modern Language Journal*, *73*(4), 395–403. https://doi.org/10.2307/326874

Shafqat, A., Arain, F., & Dahraj, M. (2020). A Corpus Analysis of Metadiscourse Markers Used in Argumentative Essays by Pakistani Undergraduate Students. *International Journal of Psychosocial Rehabilitation*, *24*, 341–351. https://doi.org/10.37200/IJPR/V24I4/PR201013

Shah, M. (2021). *Effect of Data Driven Language Learning (DDLL) on EFL Learners: A Corpus-driven Language learning Approach*. *01*, 35–55.

Shamim, F. (2017). *English as the language of development in Pakistan: Issues, challenges and possible solutions*.

Shaw, E. (2011). *Teaching Vocabulary Through Data-Driven Learning*. https://www.semanticscholar.org/paper/Teaching-Vocabulary-Through-Data-Driven-Learning-Shaw/d1758aa87bcf9211957f1f5738d4d39bd38c2f72

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, *26*(2), 184–201. https://doi.org/10.1017/S0958344014000081

Smith, M. S. (1993). Input Enhancement in Instructed SLA: Theoretical Bases. *Studies in Second Language Acquisition*, *15*(2), 165–179. https://doi.org/10.1017/S0272263100011943

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Harvard University Press Cambridge, MA.

Stenstrom, A.-B. (1994). *Introduction to Spoken Interaction, An*. Routledge. https://www.taylorfrancis.com/books/mono/10.4324/9781315845449/introduction-spoken-interaction-anna-brita-stenstrom

Sultan, D. S., Khattak, D. N.-U.-R., & Kanwal, D. A. (2021). THE MORPHOLOGICAL PATTERNS OF ENGLISH DISCOURSE MARKERS IN PASHTU SPEECH IN PAKISTAN. *PalArch's Journal of Archaeology of Egypt / Egyptology*, *18*(18).

Sun, W. (2013). The Importance of Discourse Markers in English Learning and Teaching. *Theory and Practice in Language Studies*, *3*(11), 2136–2140. https://doi.org/10.4304/tpls.3.11.2136-2140

Therefore. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/therefore

THEREFORE definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/therefore

Though. (2024, April 17). https://dictionary.cambridge.org/dictionary/english/though

THOUGH definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/though

Uchibori, A. (2013). *Comparing Computer-based and Paper-based DDL in the Beginner Level English Classroom*. https://www.academia.edu/76811162/Comparing_Computer_based_and_Paper_based_DDL_in_the_Beginner_Level_English_Classroom

WELL definition and meaning | Collins English Dictionary. (2024, April 22). https://www.collinsdictionary.com/dictionary/english/well

Yanagisawa, A., & Webb, S. (2022). Involvement load hypothesis plus: Creating an improved predictive model of incidental vocabulary learning. *Studies in Second Language Acquisition*, *44*(5), 1279–1308.

Yao, G. (2019). Vocabulary learning through data-driven learning in the context of Spanish as a foreign language. *Research in Corpus Linguistics*, *7*, 18–46. https://doi.org/10.32714/ricl.07.02

Yasmin, T., Rashid, A., & Asghar, S. A. (2021). Efficacy of Explicit Instruction of Discourse Connectives for Verbal Communication. *Review of Education, Administration & Law*, *4*(3), Article 3. https://doi.org/10.47067/real.v4i3.178

Ying, S. (2007). *An Analysis of Discourse Markers Used by Non-native English Learners: Its Implication for Teaching English as a Foreign Language*.

Yılmaz, E., & Soruc, A. (2015). The use of Concordance for Teaching Vocabulary: A Data-driven Learning Approach. *Procedia - Social and Behavioral Sciences*, *191*, 2626–2630. https://doi.org/10.1016/j.sbspro.2015.04.400

Zahra, T., & Abbas, A. (2018). Pedagogical Implications of Corpus-based Approaches to ELT in Pakistan. *Journal of Education and Educational Development*, *5*, 259. https://doi.org/10.22555/joeed.v5i2.1565

APPENDIX

Annexure A – Consent Form

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Information and Privacy Notice

The current research study is a part of thesis by Najam Irshad, a student of Mphil English Linguistics at the National University of Modern Languages (NUML), Islamabad. The study aims to test a new method of teaching called Data Driven Learning (DDL), which involves putting the learners in the driving seat of learning and makes them learn language as researchers. As a part of the study, learners will have to write three essays in total as well as complete 3 sentence making exercises. They will have to participate in total 8 lectures. There will be three groups for the study: two experimental groups, which will participate in two different versions of the DDL method, and a control group which will be taught using the traditional method of teaching. The material to be taught will include discourse markers, which are an essential part of language.

As a part of the research study, you will have to provide your name, your father's name and grade to the researcher. This privacy notice would like to inform you that your name and your scores on the essays will be strictly confidential and the scores will not be shared with any person or entity outside of the research study. The data collected from the research study will only be used for research purposes and the final results will be anonymised and aggregated, therefore it would not be possible to identify an individual from the results of the study.

Your participation in this research study will help measure the effectiveness of a new teaching method and will pave the way for improvement in teaching in the future.
Your help will be greatly appreciated!

CONSENT FORM

I _____ son/daughter of _____,
a student of class _____ at _____
confirm that I have read and understood the information provided by the researcher and agree to take part in the research study. I understand and agree that the researcher will not owe me any monetary or other benefit as a result of taking part in the research study. I shall not hold the researcher liable for any issue arising from this study.

Name: _____

Date: _____

Signature: _____

Annexure B – Key for Qualitative Analysis

Now

- To indicate present time
- To indicate time just before the present
- by this time
- with the sense of present time weakened or lost to express command, request, or admonition
- to introduce an important point or indicate a transition

So

- Result or consequence
- Inference or conclusion
- Transition or continuation
- Clarification or emphasis
- Beginning a story or providing background information

But

- In the sense of except for the fact
- to show contrast
- in the same sense as ‘yet’

Because

- In the sense of ‘for the reason that’
- To indicate causal relationship
- In place of ‘that’

Though

- In the sense of ‘in spite of the fact’
- In the sense of “in spite of the possibility”
- In the sense of “however, nevertheless”

However

- To show contrasting ideas

- To show contradicting statements
- To introduce a counterargument

Therefore

- To provide logical conclusions
- To summarize or conclude thoughts
- To show cause and effect
- To strengthen an argument

Well

- Introduction to a New Point
- Indicating Consideration or Thought
- Easing into a Strong Statement
- Clarifying or Expanding
- Softening Critique or Disagreement
- With ‘as’ in the same meaning as ‘also’

Oh

- To express surprise or realization
- To indicate agreement or acknowledgment
- To soften corrections or critiques
- To express puzzlement or confusion
- To add emphasis or intensity
- To expressing concern or dismay

Really

- To expressing surprise or seek confirmation
- To provide emphasis
- To express disbelief or doubt

I think

- To express personal opinion
- To soften statements

- To indicate modesty
- To introduce uncertainty
- To transition to a topic

Also

- To adding information or continue a thought
- To show similarity
- To emphasize equivalence
- To support a main point

Then

- To showing sequence
- To indicate cause and effect
- To transition between ideas
- To indicating a point in time

Instead

- To presenting an alternative choice
- To replace one thing with another
- To introduce a contrasting idea
- To offer a counterpoint

Perhaps

- To express uncertainty
- To suggest a possibility
- To soften a direct statement

Annexure C – Pages for Essay

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Pre-test

Essay

Name: _____

Annexure D – Pages for Sentence Making Exercise

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Pre-test

Sentence Making Exercise

Name: _____

Please make 3 sentences with each of the following words.
1. Now
2. So
3. But
4. Because
5. Though
6. However
7. Therefore

8. Well
9. Oh
10. Really
11. I think
12. Also
13. Then
14. Instead
15. Perhaps

Annexure E – Handouts

Also

1. thanks to this officer that not only the forest but **also** the happiness of the man was protected.
2. The people of Italy were **also** the inheritors of the skills and knowledge of the old Roman Empire.
3. In 1893 the antitoxin which cures diphtheria was discovered, and **also** the protective treatment for cholera.
4. In the crowd, there was **also** a loudmouthed old hag who seemed to have other
5. He conducted business, wrote poetry and songs and was **also** a patron of Michelangelo and other artists.
6. bacteria set Lister wondering whether inflammation was not **also** a type of fermentation due to bacteria getting into
7. it is obvious that their off-springs will **also** be thin, weak and unhealthy.
8. Had you **also** been blessed with a daughter then you would understand
9. our ways of living changed, but people themselves have **also** been changed.
10. to say how far the digging may extend. Work is **also** to be done at Brahmanabad, about forty-three miles
11. which gives the aerobic bacteria their chance. They **also** are always present, naturally, in sewage, and they continue
12. of cultivating them so that they were weakened, and **also** at the fact that such germs inoculated into a
13. When the uneasiness lasted longer the king **also** became displeased. In that boat there happened to be
14. has not only been reduced by famine and disease but **also** by war. We have the power to abolish war
15. it was soon found that a chemical that destroyed germs **also** destroyed the cells of the body. Injecting carbolic acid
16. However, along with the excitement of freedom, students may **also** feel stress and pressure of parental expectations to perform
17. She fought, with clenched teeth, desperately! But now I also had grown furious - at a child.

Because

1. Someone had to keep the traffic in order. Today, **because** of the big snow, there weren't many people
2. This was urgently necessary for two reasons: first, **because** educated speech under the Ottoman Empire had been a
3. In rural areas it is twice of this number **because** of the non-availability of qualified healthcare staff
4. I don't want to ask for too much, **because** I shan't get it, and I don't
5. Be good to me, because I sold it for you.
6. I was frightened, too, sometimes, **because** I thought I might die and never know
7. This was the more remarkable, **because** I was found unable to answer a single question
8. I think they ought to be very severely punished, **because** the entire system of control and regulation of foodstuffs
9. He does not flee to the cities **because** the industrial road passes through a technically-oriented agriculture
10. it has never completely eliminated the disease, apparently **because** the silkworm is not the only insect which harbours
11. War tetanus (lock-jaw) was tremendously common amongst our wounded **because** the soil of Belgium and Norther France is full
12. it has been demonstrated that ideas are not necessarily true **because** they have been believed for a long time.
13. As a matter of fact, women need good nutritious food **because** they have to feed and bring up children.
14. They may replace foreign machines **because** they are superior to foreign equipment
15. He had got rid of the sheep **because** they threatened his young trees.
16. Joel takes his learning seriously **because** his goal is to become a farmer.
17. would cease to be Muslims, not in the religious sense, **because** that is the personal faith of each individual, but
18. But above all else, we find the universe frightening **because** we cannot find any sign that life like our

But

1. No one can do this but the boy himself. All that the rest of us
2. A tyrant does not remain in the world, but the curse on him abides forever!
3. and out of these researches, which but for Fleming would not have been started came new
4. Norma almost hung up **but** restrained herself.
5. “It is not we who are stupid," answered the camel-men, "**but** you.”
6. he shocked his hosts by saying, "Wonderful, **but** penicillin could never have been discovered in a lab
7. I cannot tell you; **but** the fact remains that I have not got it.
8. All the crops finished. Nothing left," he said. **But** the gongs were still beating, the men still shouting,
9. Greek Royalist officer- Ioanne Metacas - protested strongly to his Government, **but** the invasion continued.
10. The Chinese may have introduced the smile policy, **but** the Occidentals certainly launched the warm hug.
11. There was no answer **but** the racing hiss of wind through the stiff grass.
12. Thus he tried adding antiseptics to the wine, **but** the result was unsatisfactory, so he proceeded to try
13. spinal cord introduced into dogs rendered them immune to hydrophobia, **but** the treatment was not tried on human beings till 1885
14. There were some whom he found suitable. **But** the trouble was that everyone in the community knew
15. Chinese social experiment does not only concern Asia however, **but** the West too
16. The sharp icy gusts of wind struck like steel, **but** the young man seemed to be immune to them.
17. There was water there all right- **but** it was a good 46 metres below
18. The wire was very thin, little thicker than a human hair, **but** it was about 305 meters long.
19. This was the end of his main journey; **but** it was by no means the end of his

20. It was like a mixture of mud and sulphur- **but** it was drinkable, and it would save him from…

However

1. Pasteur was devoured by fears and anxiety about the results. However, the boy was absolutely cured
2. inhabitants were driven to import fresh eggs each year. Soon, however, the disease spread to neighbouring countries
3. If, however, the patient has money problems, then his company steps in
4. “I am not at liberty to tell you that," he said. "However, I assure you, the organization is of international scope." "
5. “The fact remains that I have not got it. Everything else, however, I have got."
6. Western women have not yet succeeded in obtaining all this. "However, we Chinese are working
7. In Pasteur's case, however, we have an exception to this state of things
8. In “1763," however, an Italian priest, the Abbe Spallanzani, repeated the experiments,
9. On the appointed day, Pasteur appeared loaded with apparatus. His opponents, however, had none
10. To all who are detected in any deliberately furtive act, however harmless.
11. He gave aid on a large scale and tried to help. However, imported food could not solve the problem.
12. What it has achieved, however, is the keeping of the disease well in check
13. The real glory of this period, however, lies in fields other than political.
14. and one walked home from school on a Friday afternoon (however much homework had been assigned for the Monday)
15. However, Pasteur's work on fermentation did not stop.
16. But his voice remained resonant. However, sometimes that too quivered.
17. The Headmaster, Mr. Welldon, however, took a broadminded view of my Latin prose.
18. The marriage, however, was extremely happy, and the wife seems to have…

19. Even her expression hadn't changed. Her breaths, however, were coming faster and faster. Then the battle began.

I Think

1. Norma said. Arthur looked astounded, "Are you saying what **I think** you are?"
2. "In what way?" His voice was guarded. "**I think** you felt" -she gestured again- "that I was only…
3. Well - yes, **I think** you might put it like that.
4. Pardon me, good lady. Had you been there, **I think** you would have begged the ring of me
5. to get the money to buy the combs. And now **I think** we should have our dinner.
6. Britain and they are all members of the Nation. Now **I think** we should keep that in front of us
7. “However, I assure you, the organization is of international scope." "**I think** you'd better leave, "Arthur said
8. that will destroy a whole city in a few minutes, **I think** you're ... you're mad-wickedly, horribly mad.
9. had seen so much and knew so much And then **I think** I felt prouder of him than ever before,
10. could hardly avoid noticing the tear stains on her cheeks. "I think I'll have an apple," she said.
11. Norma slid beneath the covers. "Well, **I think** it's intriguing," she said.
12. I do not say that other countries are free from it, but, **I think** our condition is much worse
13. I think that I had better sharpen my pencil
14. when they indulge in black-marketing, **I think** they ought to be very severely punished
15. MRS. MELDON: What is your invention, Henry? CORRIE: Oh, **I think** we'd better not talk about it!

16. "All that talk about the button," Norma said." **I think** you-well, misunderstood me."
17. That's why **I think** your cleverness can only destroy it
18. He provides, **I think**, the outstanding example of a scientist whose work
19. not only we ourselves are wondering but, **I think**, the whole world is wondering at this unprecedented…
20. A citizen who does black-marketing commits, **I think**, a greater crime than the biggest and most
21. Pasteur is, **I think**, unique amongst scientists in the versatility of his…
22. I'll grow into it It... it looks like her, **I think**."

Instead

1. "I'll play some of the old tunes." But **instead** of the harmonica, he brought out the blanket.
2. Everything was changed. Even the air. **Instead** of the harsh dry winds that used to attack
3. on behalf of France, she might have proved the winner **instead** of the loser.
4. infection in surgical cases had become the exception **instead** of the rule; now it was the other way
5. which produce light, to occur in one direction only **instead** of in all directions at right angles…
6. been largely replaced by what was called the aseptic method. **Instead** of chemicals heat was used to sterilize instruments…
7. but, instead of recovering like your camel, she died, and **instead** of getting a fee I was compelled to dig
8. Books were printed in people's own languages **instead** of Latin and this encouraged more people to learn
9. in the family. At night he consulted his wife. But **instead** of replying, she silently lifted a corner of the…
10. Then I realized that I had written fifty-six **instead** of six.
11. They ought to be horse-knackers **instead** of soldiers. And tin hats, too!
12. "I'd give anything to be ploughing **instead**!"
13. "No, you're lying! The rockets will come back!" **Instead**, he stroked Laura's head against him and said…

14. Now he had only four sheep but, **instead**, a hundred beehives.
15. does not want to be urbanized, but is **instead** kept on the spot and incorporated in local small
16. An idea came to him. **Instead** of listening to Shamim Ahmed's request on his threshold, he
17. Nobody believed his story of the string. **Instead** people laughed at him.
18. But **instead** she got up and stretched and wrinkled her nose.

Now

1. And **now** I think we should have our dinner.
2. It is perhaps a rather appropriate time just **now** for refreshing one's memory of Pasteur's work.
3. CORRIE: **Now**, now, now, Charlotte, not again please.
4. She had been on the defensive before but **now** she attacked.
5. And **now** they were hers, but her hair was gone.
6. cleared away, dilapidated walls torn down and five houses restored. **Now** there were twenty eight inhabitants, four of them young
7. **Now** there was not a single scrap of food left.
8. We have broken the shackles of slavery; we are **now** a free people.
9. We come **now** to a new departure in Pasteur's scientific work
10. **Now**, the idea that life could originate casually and spontaneously…
11. His daughters' ill-usage had really made him go mad. And **now** the loyalty of this worthy Earl of Kent showed
12. What else? His sleeping-bag? No, for it was **now** the winter season, which meant that the nights were long
13. The numbers doubled by 1650 and by 1850 doubled again to 1,000 million. **Now** the world population is over 3,000 million.
14. surgical cases had become the exception instead of the rule; **now** it was the other way about again.
15. "Oh, Yes." Norma repressed a smile. She was sure **now** it was a sales pitch.
16. After a few minutes he got up. By **now** it was past seven.
17. This is the time to make real the promise of democracy. **Now** it the time to rise from the dark

18. "That book of yours," I said, "I suppose, **now**, it contains all the good actions that we men
19. I whispered. "I propose to deposit fifty-six dollars **now** and fifty dollars a month regularly."
20. Remember that you are **now** a sovereign legislative body
21. and roses, leeks and snapdragons, celery and anemones. It was **now** a village where one would like to live.
22. Margaret was watching the hills. **Now** there was a long, low cloud advancing, rust colour
23. The Roman Catholics and the Protestants persecuted each other. Even **now** there are some States in existence where there are
24. not really believed his Dad would send Granddad away. But **now** - there it was - the going-away gift.
25. said, winking at me, "I have cheated death many times! **Now**, I've reached the years the Good Book allows

Oh

1. "What?" "How will you have it?" "**Oh**".... I caught his meaning and answered without even trying
2. CORRIE: **Oh**, I don't know. It's fantastic thought, that,
3. I thought then, Henry, I just can't tell you, but **oh**! I prayed for him, Henry
4. "**Oh**, my dear love," said Portia, "despatch all business and
5. There was no decency in his death! . . . **Oh**, my God! My God!
6. CORRIE (almost forgetting his grievance). **Oh**, yes! I'd forgotten that!
7. They are very beautiful, aren't they, Henry. CORRIE: **Oh**, Yes! Quite nice!
8. until this night when he brought home the blanket. "**Oh**, yes, it's a fine blanket," Peter said
9. Give it to me quickly," said Della. **Oh**, and the next two hours seemed to fly.
10. That seems to me a horrible thing! CORRIE. **Oh**, a mother's feelings, of course.
11. young man said to the carpet dealer. "With pleasure, sir." "**Oh**, don't bother to take it down.
12. A rocket, Harry? To go back to all that trouble? **Oh**, Harry!" "But you must want to go back.

13. Is that what you wanted to show me?" I asked. "**Oh**, no, no," he said as he began to climb
14. CORRIE: Oh, now, don't be bitter. Charlotte
15. “Not so fast, Jim!" And then she jumped up and cried, "**Oh**, **oh**!" Jim had not yet seen his beautiful gift.
16. The young man frowned as if to suggest, "**Oh** so much." "You have only to select, sir,"
17. but in this attempt he dropped it. "**Oh**, sorry," he exclaimed and bending down picked it up.
18. was the news which had so distressed him, he said: “**Oh**, sweet Portia, here are a few of the unpleasantest…
19. "Something wrong?" Mr. Steward introduced himself. "**Oh**, the-" Arthur pointed toward the living room and smiled.
20. He had promised that he would explain this to me later. "**Oh**, yes," he said.
21. Mrs. Lewis?" he inquired politely. "Yes?" "I'm Mr. Steward." "**Oh**, Yes." Norma repressed a smile.
22. Then suddenly as if remembering something he called out. "**Oh** yes... Could you let me have a list of…
23. It was amazing that Minta could laugh and say, "**Oh**, you... " the way she did year after year.
24. It's not an Earthman's house any more." "**Oh**, your imagination!" He put on his coat and tie. "

Perhaps

1. It is **perhaps** a rather appropriate time just now for refreshing one'
2. We'll utterly obliterate whole cities, **perhaps** a whole nation.
3. she was screaming in wild hysterical shrieks. **Perhaps** I should have desisted and come back in
4. Chances are a plague did this town in, sir." "**Perhaps**. I suppose this is one of those mysteries we
5. Economic decentralization, which is **perhaps** the most important step in Chinese domestic policy
6. **Perhaps** the real shift will come when men fully realize,
7. But I am a proud man, Hannah. **Perhaps** you'll understand that.

8. **Perhaps** an enemy of this country might make a similar
9. till the sun goes down, they'll settle somewhere else **perhaps**...."
10. He supposed it was community property, or **perhaps** belonged to people who cared nothing about it.
11. Of Buddhist memorials, **perhaps** enough has been said, although it deserves to be
12. He was **perhaps** even more of a patriot than of a scientist
13. The precipice was sheer, and the top of it was **perhaps** farther away than the sea beneath him.
14. The thought occurred to me: "**Perhaps** he suspects I have swag in this sack."
15. Next, **perhaps** in fame but certainly not inferior in importance, is
16. Or **perhaps** it was that he saw no need for it.
17. Salah is fighting a desperate battle for survival, and **perhaps** losing the contest.
18. Alas, his little cane, which was **perhaps** lost at the time of the accident, was not
19. **Perhaps** they will advise me.
20. **Perhaps** this is one of the best examples of a
21. Above the yams were, **perhaps**, three dozen rows of potatoes.
22. There was silence for a while. Then Maulvi Abul said, **perhaps** to reassure himself more than his wife: "Don't
23. the persons he had observed that evening. He had been, **perhaps**, too deeply engrossed in himself.
24. books, poor innocent books, you are lying there still: covered, **perhaps**, with mud by this time,
25. lie there in a living death, and your fate is **perhaps** worse than you deserved.

Really

1. a germ which Pasteur has described as a micrococcus was **really** a bacillus (the difference is merely one of shape).
2. It is **really** a delightful sensation.
3. CORRIE: I'd like to, of course, but I **really** must finish up these things.
4. "I **really** must get at my Christmas shopping," she mentioned
5. But Peter had not **really** believed Dad would do it... until this night
6. **Really**, Charlotte, you're insufferable! You're absolutely insufferable!
7. Yes, I know, Charlotte, but you **really** ought not to dwell too much on your sorrow.
8. He was afraid I'd wake his baby sister. "You **really** take good care of sister," I commended.
9. You don't **really** think they'd kill somebody, do you?
10. Do you **really** think you knew your husband?"
11. They ought **really** to give me a title too. Supposing I say
12. I did not **really** doubt his care and conscientiousness, but
13. One wonders how much of our gratitude to him is **really** due to her.
14. Now he was able to attend lectures by **really** great chemists, such as Balard, the discoverer of bromine,
15. Many boys attempt seriously to make good, and **really** have the native ability to do so,
16. Thirdly, learn to judge who are really true and **really** honest and unselfish servants of the State
17. I come in. I'm going to make war horrible, **really** horrible!
18. What is it like to be **really** hungry?
19. These black marketeers are **really** knowing, intelligent and ordinarily responsible people,
20. She was like her mother. Inside, where it **really** mattered she was like her mother
21. that has not been settled yet- but it's not **really** my job.
22. The Angel said I **really** need not be alarmed, that everything had been noted,
23. The position was becoming **really** serious for the Allied Army of Occupation stationed here

24. for it is here that the desert **really** starts.
25. height I know and you all know what has been **really**. the result of this.
26. Thirdly, learn to judge who are **really** true and really honest and unselfish servants of the state
27. Until he does this, he does not **really** try, although he thinks he's trying
28. Antonio signed the bond, thinking it **really** was (as the Jew said) merely in sport

So

1. there were barely enough to go round, **so** that the houses of the foreigners were ransacked and
2. However, we Chinese are working **so** that the women of the world can be equally
3. I often sit here so he can see me. " Ah, **so** that was it…
4. he said as he pushed down the top wire **so** he could cross the fence.
5. They had to be pleased first. **So** he declared slowly, "Tonight, all my children will get
6. "I'll not be seeing you in the morning, **so** I came over to say good-bye."
7. There were no more diseases after zymosis, **so** I concluded there was nothing else the matter with
8. “You tell that to the Marines, my son!" **So** far I do not know how long I strayed
9. "I am terribly sorry, Maulvi ji. My wife is unwell. **So** the maid has cooked the food." He, sincerely considered
10. No. We're altogether too pugnacious, we human beings. **So** the only thing to do, then, is to make
11. As most of the population is illiterate and conservative, **so** the parents do not send their daughters to school.
12. 'Anything?' asked Terbut. 'Anything,' Jorkens replied, '**so** long as he sticks to it
13. he requested that the learned young Doctor Balthasar (**so** he called Portia) might be permitted to plead in
14. —and **so** he created a new branch of Science

Then

1. he had not eaten since the previous nightfall. **Then** he had found a dried piece of mackerel's tail
2. he had boiled it: and even **then** he had to hold his nose while he drank
3. he clung the stern with both of his hands. **Then** he sat down and remained quiet.
4. Which one?" he asked, laughing. **Then** he sat down on a big oak stump
5. you drawing it all out again?" he asked in surprise. **Then** I realized that I had written fifty-six instead
6. I hugged him tightly and pressed my cheek against his. **Then** I rose, buttoned my coat, pulled my hat down
7. He'd seen so much and knew so much. And **then** I think I felt prouder of him than ever
8. “Will you run?" asked the woman. "Yes'm," said the boy. "**Then** I won't turn you loose," said the woman.
9. **Then** I wondered how long I had to live.
10. I would make them all learn English: and **then** I would let the clever ones learn Latin as
11. It almost made itself into a dress for her. And **then** she put it up on her head again, nervously
12. He had lost his only son, **then** his wife. He had withdrawn into this solitude
13. I sat for a while frozen with horror; and **then**, in the listlessness of despair, I again turned over
14. That'll break the ice a bit. **Then** when she realises how important my work is, and
15. No, I can take care of myself. But **then**, when they hear what my invention is, they'll
16. your lord will never more break his faith with you.” “**Then** you shall be his surety,” said Portia.
17. This is it. I've tried all kinds of land!" **Then** he smelled the dirt. He whiffed and whiffed the
18. **Then** he tried the effect of dissolving the crystals and

19. At been at Oxford long when the War began, and **then** he went off and enlisted.
20. Since **then** I bank no more. I keep my money in
21. I'm back at school again until 3 o'clock. **Then** I go back home and work.
22. **Then** he realized that it was covered with tiny stones.

Therefore

1. would never take any interest for the money he lent; **therefore** there was great enmity between this covetous Jew and
2. a gigantic task, for state education was unknown in Turkey. **Therefore** there were two problems: to teach the masses and
3. none of us acting alone can achieve success. We must **therefore** act together as a united people, for national reconciliation,
4. who must have died shortly in any case. Let us **therefore** compel the wretch to dig her grave, and then
5. to the second and third millennia B-C. and is, **therefore**, contemporary with the prehistoric sites of Mesopotamia, with which
6. and your life lies at the mercy of the duke; **therefore**, down on your knees and ask him to pardon
7. The population is **therefore** growing at the rate of 6.6 per 1,000 of the population.
8. Now no one would have bought these books. I **therefore** had to throw them away or wipe them off
9. My young friends, I would, **therefore**, like to tell you a few points about which
10. Parental interference and pressure is, **therefore**, one of the most significant external trigger for pressures
11. It, **therefore**, places on you the gravest responsibility as to how
12. The Jew shall have nothing but the penalty. **Therefore** prepare, Shylock, to cut off the flesh; but mind
13. place today in the world's most populous country is **therefore** the education and re-education of man.
14. considerable distance from the well to the home. It was, **therefore**, used very sparingly for bathing and cleaning purposes.

15. **Therefore**, we must learn a lesson from this.

Though

1. Portia hearing this, **though** the kind-hearted lady was not at all offended
2. those who sat inside, for they were listening in silence **though** the music was foreign.
3. **Though** the watch was so fine, it never had a
4. Even though it was a snowy, windy night, even **though** the worker, his wife and children couldn't see
5. with a slightly sad expression towards the royal platform, as **though** he asked some question with his eyes.
6. by his witty sayings would keep up his good-humor, **though** he could not refrain sometimes from jeering at his
7. His camel fell heavily, knocking him off its back; and **though** it was not really injured, it was so shocked
8. This treatment, **though** it was not at once adopted, was very successful
9. She watched his face as **though** it were the face of a stranger.
10. where it really mattered she was like her mother, even **though** she had her father's dark eyes
11. she borrowed it back again. She didn't mean it **though**... she loved having Minta borrow her things.
12. Prices have risen so suddenly that it seems as **though** all the cows and buffaloes in the country have
13. Director of Science at the ficole Normale in Paris. This, **though** an advancement, was not an unmixed blessing
14. people today believe that diseases are caused by evil spirits. **Though** astrology and fortune-telling are still practised, they do
15. By the advice of the physicians, Cordelia, **though** earnestly desirous of seeing her father, was prevailed upon
16. and I don't want to ask for too little, **though** I'll probably get that anyhow.
17. More than the bother, **though**, Joel hates the ugliness.
18. animals need - corn, oats, and hay. They sell the surplus, **though** most of their money comes from selling the animals

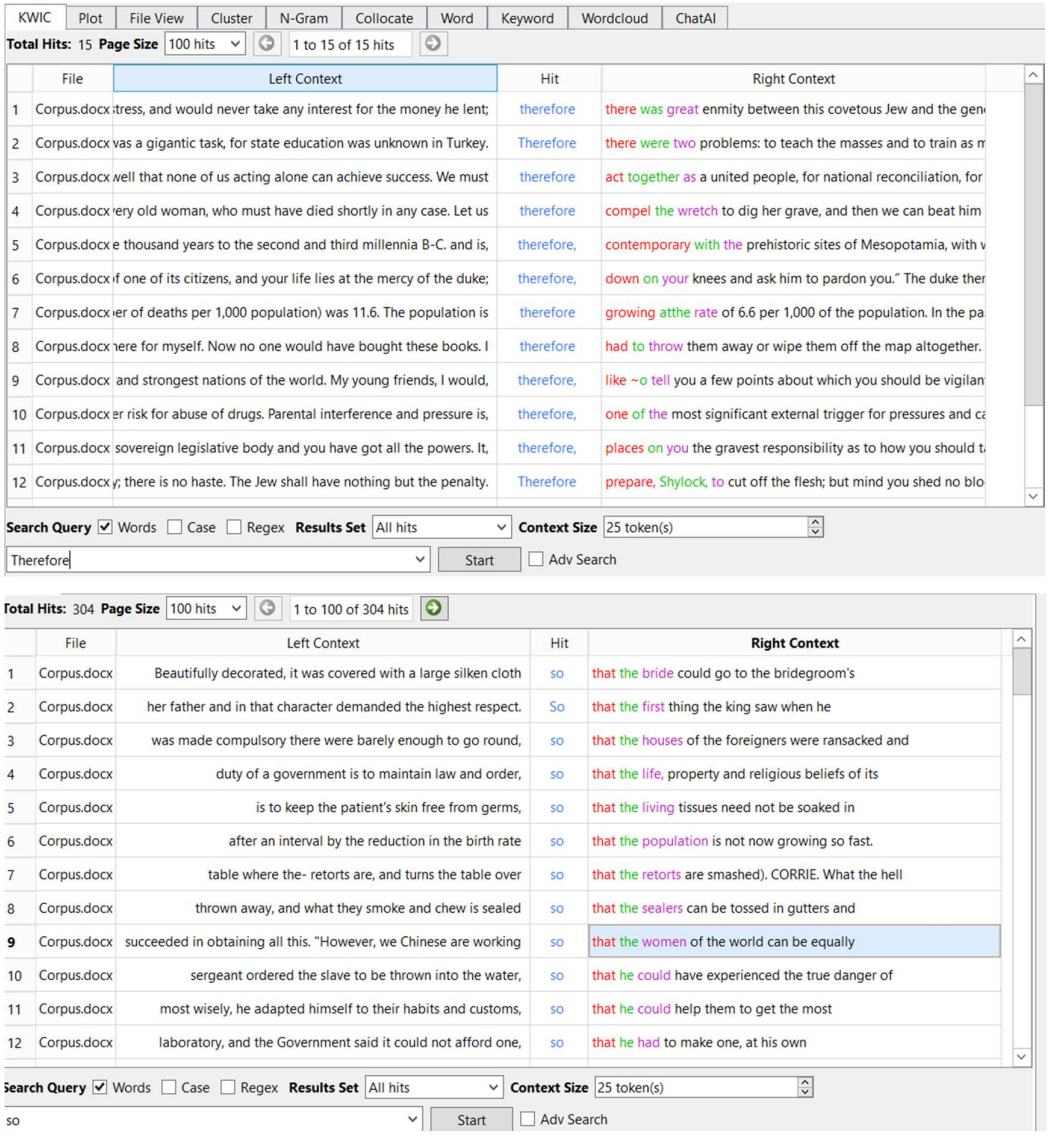
19. a comparatively low temperature was enough to kill harmful bacteria, **though** not all
20. the result, viz., that the birds still resisted the disease, **though** others, which had not been previously dosed with the
21. ill at all and played about the laboratory very happily, **though** Pasteur was devoured by fears and anxiety about the
22. essential point of the matter is that the agricultural labourer, **though** possibly deprived of farming machinery, must not and does
23. Joel is a good student, though reluctant to discuss it.
24. The Tuaregs, **though** their life is primitive, are a people of great
25. In fact she knocked my glasses flying and they fell, **though** unbroken, several feet away from me on the kitchen

Well

1. soaked in carbolic, which tends to destroy the tissue as **well** as the germ. Hence, though the instruments and the
2. the sea was the safest way to trade, as **well** as the quickest. Overland routes were dangerous, with their
3. This situation is quite grave and gloomy. The quality as **well** as quantity of food for female and male is
4. but I didn't snatch people's pocketbooks. **Well**, I wasn't going to say that." Pause. Silence. "
5. I was only thinking of myself." "Oh." "I wasn't." "Norma-" "**Well**, I wasn't.
6. The Oxford team included trained chemists as **well** as bacteriologists, and had all the equipment that Fleming…
7. It was an area notorious for sandstorms as **well** as for dried-up waterholes. Christopher soon experienced one
8. Black-marketing is another curse. **Well**, I know that blackmarketeers are frequently caught and punished.
9. Norma slid beneath the covers. " **Well**, I think it's intriguing,"
10. I want a pair of blue suede shoes," said the boy. " **Well**, you didn't have to snatch my pocketbook
11. I've ever had. MRS. MELDON. Oh, yes. CORRIE. (dashed) Well, you don't seem very excited about it. MRS .

12. The house looked pitiful and decrepit to him. "**Well**, it's Arshak's house, too."
13. … was introduced later on and it went up progressively. **Well**, the whole principle was to create a mentality
14. Navasard repeated the words to himself. "**Well**, then, that means he'll soon be home."
15. How much time do you spend on homework at home?" "**Well** actually I do my homework at school.
16. The boy was back in no time. "**Well**? Did you see Arshak?" "What's he doing?

Annexure F – Samples of Activities



KWIC | Plot | File View | Cluster | N-Gram | Collocate | Word | Keyword | Wordcloud | ChatAI

Total Hits: 15 **Page Size** 100 hits 1 to 15 of 15 hits

	File	Left Context	Hit	Right Context
1	Corpus.docx	tress, and would never take any interest for the money he lent;	therefore	there was great enmity between this covetous Jew and the gen
2	Corpus.docx	vas a gigantic task, for state education was unknown in Turkey.	Therefore	there were two problems: to teach the masses and to train as m
3	Corpus.docx	well that none of us acting alone can achieve success. We must	therefore	act together as a united people, for national reconciliation, for
4	Corpus.docx	ery old woman, who must have died shortly in any case. Let us	therefore	compel the wretch to dig her grave, and then we can beat him
5	Corpus.docx	e thousand years to the second and third millennia B-C. and is,	therefore,	contemporary with the prehistoric sites of Mesopotamia, with v
6	Corpus.docx	f one of its citizens, and your life lies at the mercy of the duke;	therefore,	down on your knees and ask him to pardon you." The duke ther
7	Corpus.docx	er of deaths per 1,000 population) was 11.6. The population is	therefore	growing atthe rate of 6.6 per 1,000 of the population. In the pa
8	Corpus.docx	here for myself. Now no one would have bought these books. I	therefore	had to throw them away or wipe them off the map altogether.
9	Corpus.docx	and strongest nations of the world. My young friends, I would,	therefore,	like ~o tell you a few points about which you should be vigilan
10	Corpus.docx	er risk for abuse of drugs. Parental interference and pressure is,	therefore,	one of the most significant external trigger for pressures and ca
11	Corpus.docx	sovereign legislative body and you have got all the powers. It,	therefore,	places on you the gravest responsibility as to how you should t
12	Corpus.docx	y; there is no haste. The Jew shall have nothing but the penalty.	Therefore	prepare, Shylock, to cut off the flesh; but mind you shed no blo

Search Query ☑ Words ☐ Case ☐ Regex **Results Set** All hits **Context Size** 25 token(s)

Therefore Start ☐ Adv Search

Total Hits: 304 **Page Size** 100 hits 1 to 100 of 304 hits

	File	Left Context	Hit	Right Context
1	Corpus.docx	Beautifully decorated, it was covered with a large silken cloth	so	that the bride could go to the bridegroom's
2	Corpus.docx	her father and in that character demanded the highest respect.	So	that the first thing the king saw when he
3	Corpus.docx	was made compulsory there were barely enough to go round,	so	that the houses of the foreigners were ransacked and
4	Corpus.docx	duty of a government is to maintain law and order,	so	that the life, property and religious beliefs of its
5	Corpus.docx	is to keep the patient's skin free from germs,	so	that the living tissues need not be soaked in
6	Corpus.docx	after an interval by the reduction in the birth rate	so	that the population is not now growing so fast.
7	Corpus.docx	table where the- retorts are, and turns the table over	so	that the retorts are smashed). CORRIE. What the hell
8	Corpus.docx	thrown away, and what they smoke and chew is sealed	so	that the sealers can be tossed in gutters and
9	Corpus.docx	succeeded in obtaining all this. "However, we Chinese are working	so	that the women of the world can be equally
10	Corpus.docx	sergeant ordered the slave to be thrown into the water,	so	that he could have experienced the true danger of
11	Corpus.docx	most wisely, he adapted himself to their habits and customs,	so	that he could help them to get the most
12	Corpus.docx	laboratory, and the Government said it could not afford one,	so	that he had to make one, at his own

Search Query ☑ Words ☐ Case ☐ Regex **Results Set** All hits **Context Size** 25 token(s)

so Start ☐ Adv Search

KWIC | Plot | File View | Cluster | N-Gram | Collocate | Word | Keyword | Wordcloud | ChatAI

Total Hits: 68 **Page Size** 100 hits 1 to 68 of 68 hits

	File	Left Context	Hit	Right Context
1	Corpus.docx	in the Indus valley region before the arrival of Aryan invaders and	also	the fact that more than one ethnic type was then present. The mea
2	Corpus.docx	ɔe happy!" It was thanks to this officer that not only the forest but	also	the happiness of the man was protected. He delegated three rang
3	Corpus.docx	made them costly in time and resources. The people of Italy were	also	the inheritors of the skills and knowledge of the old Roman Empir
4	Corpus.docx	. In 1893 the antitoxin which cures diphtheria was discovered, and	also	the protective treatment for cholera. Before the discovery of the a
5	Corpus.docx	cation; the Alpine and Mongolian types have been identified and	also	the Proto-Australiod group, typical of the aboriginal of south and
6	Corpus.docx	them by accident, and that I believe I could have worn again. And	also	the raffle I had joined for a motor car. The Angel said I really neec
7	Corpus.docx	uch money as he desired from the angels. In the crowd, there was	also	a loudmouthed old hag who seemed to have other views. In a lou
8	Corpus.docx	losophy. He conducted business, wrote poetry and songs and was	also	a patron of Michelangelo and other artists. Like Florence, he was f
9	Corpus.docx	t and his determination not to be defeated by circumstances. It is	also	a tribute to the engineers and their team work, and to their faith ir
10	Corpus.docx	ue to bacteria set Lister wondering whether inflammation was not	also	a type of fermentation due to bacteria getting into the wound. An
11	Corpus.docx	neral thing, the boy wins out in such controversies. And he should.	Also,	be it said, the parent whom I have offended usually comes around
12	Corpus.docx	ıg abuse and mental health disorders. Adequate measures should	also	be taken by the authorities in education to implement compulsory

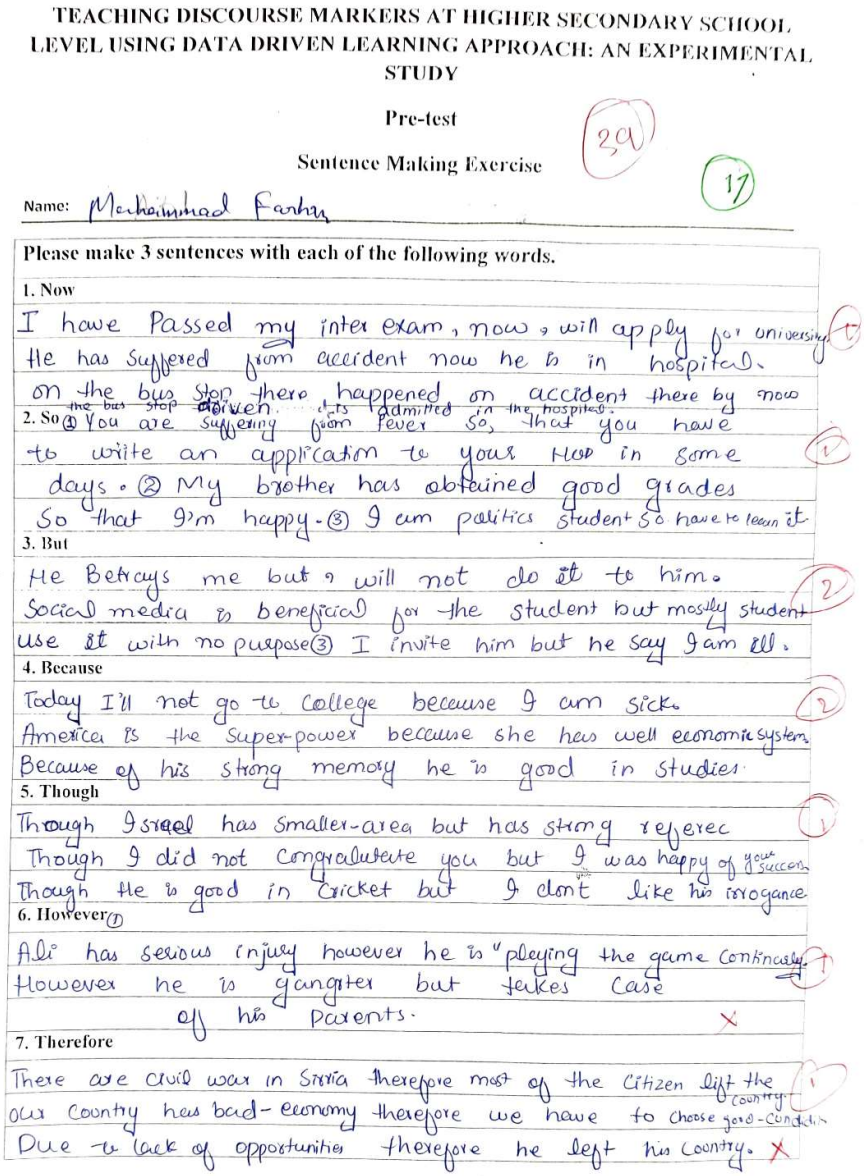
Search Query Words Case Regex **Results Set** All hits **Context Size** 25 token(s)

also Start Adv Search

1. he had not eaten since the previous nightfall. **Then** he had found a dried piece of mackerel's tail
2. he had boiled it: and even **then** he had to hold his nose while he drank
3. he clung the stern with both of his hands. **Then** he sat down and remained quiet.
4. Which one?" he asked, laughing. **Then** he sat down on a big oak stump
5. you drawing it all out again?" he asked in surprise. **Then** I realized that I had written fifty-six instead
6. I hugged him tightly and pressed my cheek against his. **Then** I rose, buttoned my coat, pulled my hat down
7. He'd seen so much and knew so much. And **then** I think I felt prouder of him than ever
8. "Will you run?" asked the woman. "Yes'm," said the boy. "**Then** I won't turn you loose," said the woman.
9. **Then** I wondered how long I had to live.
10. I would make them all learn English: and **then** I would let the clever ones learn Latin as
11. It almost made itself into a dress for her. And **then** she put it up on her head again, nervously
12. He had lost his only son, **then** his wife. He had withdrawn into this solitude
13. I sat for a while frozen with horror; and **then**, in the listlessness of despair, I again turned over
14. That'll break the ice a bit. **Then** when she realises how important my work is, and
15. No, I can take care of myself. But **then**, when they hear what my invention is, they'll
16. your lord will never more break his faith with you." "**Then** you shall be his surety," said Portia.
17. This is it. I've tried all kinds of land!" **Then** he smelled the dirt. He whiffed and whiffed the
18. **Then** he tried the effect of dissolving the crystals and
19. At been at Oxford long when the War began, and **then** he went off and enlisted.
20. Since **then** I bank no more. I keep my money in
21. I'm back at school again until 3 o'clock. **Then** I go back home and work.
22. **Then** he realized that it was covered with tiny stones.

1. "I'll play some of the old tunes." But **instead** of the harmonica, he brought out the blanket.
2. Everything was changed. Even the air. **Instead** of the harsh dry winds that used to attack
3. on behalf of France, she might have proved the winner **instead** of the loser.
4. infection in surgical cases had become the exception **instead** of the rule; now it was the other way
5. which produce light, to occur in one direction only **instead** of in all directions at right angles…
6. been largely replaced by what was called the aseptic method. **Instead** of chemicals heat was used to sterilize instruments…
7. but, instead of recovering like your camel, she died, and **instead** of getting a fee I was compelled to dig
8. Books were printed in people's own languages **instead** of Latin and this encouraged more people to learn
9. in the family. At night he consulted his wife. But **instead** of replying, she silently lifted a corner of the…
10. Then I realized that I had written fifty-six **instead** of six.
11. They ought to be horse-knackers **instead** of soldiers. And tin hats, too!
12. "I'd give anything to be ploughing **instead**!"
13. "No, you're lying! The rockets will come back!" **Instead**, he stroked Laura's head against him and said…
14. Now he had only four sheep but, **instead**, a hundred beehives.
15. does not want to be urbanized, but is **instead** kept on the spot and incorporated in local small
16. An idea came to him. **Instead** of listening to Shamim Ahmed's request on his threshold, he
17. Nobody believed his story of the string. **Instead** people laughed at him.
18. But **instead** she got up and stretched and wrinkled her nose.

Annexure G – Study Samples



TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Pre-test

Sentence Making Exercise

Name: Mehammad Farhan

Please make 3 sentences with each of the following words.

1. Now

I have Passed my inter exam, now, will apply for university.
He has suffered from accident now he is in hospital.
on the bus stop there happened on accident there by now the bus stop driven. he is admitted in the hospital.

2. So

① You are suffering from fever so, that you have to write an application to your HOD in some days. ② My brother has obtained good grades so that I'm happy. ③ I am politics student so have to learn it.

3. But

He Betrays me but, will not do it to him.
Social media is beneficial for the student but mostly student use it with no purpose ③ I invite him but he say I am ill.

4. Because

Today I'll not go to college because I am sick.
America is the super-power because she has well economic system.
Because of his strong memory he is good in studies.

5. Though

Though Israel has smaller-area but has strong referec
Though I did not Congraluteate you but I was happy of your success.
Though he is good in Cricket but I dont like his irrogance.

6. However ①

Ali has serious injury however he is "playing the game continually.
However he is gangster but takes care of his parents.

7. Therefore

There are civil war in Sirria therefore most of the Citizen left the country.
our country has bad-economy therefore we have to choose good-condition
Due to lack of opportunities therefore he left his country.

8. Well

He plays Cricket well
He is well in swimming.
He is well in everything.

9. Oh

Oh, That was a good shot.
Oh! Do you take him.
Oh, That is a good Point about his leggrey.

10. Really

I really take cause of you.
I am really sorry, I did wrong.
I am really a good Person but, I have no good friends.

11. I think

I think Ali is in depression
Think Pakistan will never recognize the Isreal as a state.
I think Pakistan's main Problem is eeonomy.

12. Also

Being a Hindko speaker I can also speak Pashto.
Being a Physical, Shrudinger also constributed in physical chemistry.
I am political science student but I am also interested in English.

13. Then

Ali was a notorious boy big then he was kia.
I always play Football then, started to play tennis.

14. Instead

Most of our teachers Prefer white-board instead of Black-board.
In our country instead of Parliamentary system, the President system is more Beneficial
I play Foot ball instead of Popular game Cricket.

15. Perhaps

If we were not separated from India perhaps we would face common situations.
If I have not choosen P.S perhapse. I would have choose chemistry.
If Ottoman empire was not break Perhaps they will be or sup or power

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Pre-test

Essay

Name:

3 1

Influence Of Social Media on Modern Society:-

Social media has become a real important Part of this modern world. Everybody uses it. Everything and everyone is influenced by it.

We can say that separation of Social media and -today's People is impossible. It has become integral Part of modern Society. It is mostly using by young People. Social media was introduced in the begining of 21st Century. Nowdays everybody Posts on it and becoming Popular over the night. I can say that everyone who use mobile uses social media and have account on different Platforms of Social media. Social media is a platform which contains other platforms to. Such as Facebook, twitter Google etc. It Provide a plateform for self-expression and Communication. It is a plate form of freedom of Speech.

It can also have negative effects on mental health and to well-being exessive social media use can lead to feeling of anxiety depression and low self stem. Additionally it is crucial for Parents, Educators and other influences to educate youth about responsible social media use and maintaining a healthy balance between online and offline life. we can get wrong information from social media. Through different Apps people easily became victims of online fraud. Due to over use of social media kias are unable to concentrate on their education.

Social media makes easier our work. It makes businesses more online and Profitable. Now everybody has understand that computer skills are very importence to learn. Due to it we can advertise our Product and reach the People. A lot of people make money through vlogs, reels and Fiktox. There are no sensorship on it.

In the suggestion I can suggest that everybody have to get benifits from it the way they can get. Many websites are available for education.

Purpose - Entertaining stuffs are also available on it. People should not do fraud on it Social media should be a Social loving media.

EACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Post-test

Sentence Making Exercise

41 22

Name:

Please make 3 sentences with each of the following words.

1. Now

I am afraid I have to go now.
You must be very hungry now.
Now youre coming to work with us.

2. So

She was unhappy, so she left her husband.
I had no idea you were so dedicated.
why are you looking so sad?

3. But

My computer is broken but I dont have money to fix it.
I em very tired, I need to sleep but I can not sleep.
She worked hard for the exam but failed.

4. Because

The girl ran because she was afraid
I was not very happy because of his illness.
I left the house because I was bored.

5. Though

He failed in the exam, even through he studied hard.
we are still very hungry, even though we ate very much
Even though they were million aries, they live old house.

6. However

I don't like him, however I agree that he is a good manager.
I really enjoyed that movie. However I Prefer the book
My friends wanted to go outside. However it is raining.

7. Therefore

She came fast. Therefore she got a good seat.
I am sleepy. Therefore I'm going to bed.
I was tired therefore I fell a sleep.

8. Well

The book is lively and well written. X
Alex Speaks French as well as me. ✓ (1)
I know very ~~well~~ where to be silent. X

9. Oh

oh no! The Park was crowded.
oh No! He is not coming to the Party! (1)
oh my God, I missed the bus.

10. Really

I think I am a really hands-on man (1)
I really need to talk to the doctor.
This Cell Phone is really expensive.

11. I think

I think she is a very good
I think you are misbehaving about this topic. (2)
I think my sister was luck woman.

12. Also

My mother also works very hard.
She Speaks French and also Speaks English. (2)
Her husband also wanted custody of the children.

13. Then

Everything since then has been a bonus. (2)
You wash your hand first and then eat.
Download the file first, than install it.

14. Instead

We shake hands instead of bowling.
Listen to what you know instead of what you fear. (1)
Henery went there instead of her mother.

15. Perhaps

I think Perhaps we want more Concious life.
Perhaps Abraham Can solve this Problem. (2)
Perhaps nature is our best assurance of immortality.

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Post-test

Essay

11 8

Name:

PROS AND CONS OF INTERNET:-

PROSE OF INTERNET:-

The role of the internet in the modern world cannot be understand. Nowday every Person use The internet -le do their daily tasks. People in different fields liks officus, Coneges, hospitel etc. use their electronic devices like laptops, Computer, cell Phones etc. te make their work fast. The internet has also made access te information easier we can learn about the whole universe with just a single click by using the internet we can easily communicate and share information with other people around the world with the help of email instant messaging video calls etc. The internet delivers a wide Variety of advantges. It not only enable People te share information but also serves as a place te share information and media digitally. This feature has benefilled the fields of education end research the most Many services are now Provided on the internet such as online booking, hotel reservation etc. The internet has made everything a lot more accessible and quik. Most Organization around the world advertese their Vaconcies on he internent So

People can search for different types of jobs around the world. The internet Provide different types of entertainment to People, bad it music, movies, theater, entertainment to to People, live matches. or live broadcasts. It is difficult to name all of the benefits and advantages of the internent. This is because the internent has become so entangled and integratad in to our daily lives that it has an influence on everything we experience around us.

CONS OF INTERNET:—

While the internet Provides us with all tools Products and services we need right at our doorstep. at the same time, it isolates us from the world outside. As we get more occustomed to ordering everything online, be it clothes, food drinks, grocery, Commodities, or even Paying bills, getting out of the houses has become less frequent. This has caused health issues and various mental health issues such as social anxiety, insomnia and kinds are the must influence by the internet because they are the generation which has seen the immense use of the internet. They are moulded to a life dependent on the internet. This hindes their learning capabilities and real life Problem-Solving skills because they are accustomed to using their mobile for every tasks.

Hence, we have learned that why the internet is vast as it is, it will always have a Positive side and negative are - As the internet evolves and All Continous to take hald an the web, we will learn to about more Positive and negative.
However untill then we must use Caution.

TEACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL LEVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Delayed Post Test

Sentence Making Exercise

42

24

Name:

Please make 3 sentences with each of the following words.

1. Now

Now lets try and locate your phone.
Now I'm sure you see it.
Now you won't forget me!

2

2. So

She likes computers, so she bought one
So she still wanted to hate her.
She looked cute. So I gave her a little toffee.

3

3. But

we have some cold day's, but mostly it is warm.
But it is not more comfortable.
I know but Sara is sick.

2

4. Because

Don't let him suffer because of me.
I like you very dearly because you are my friend.
She come because of Sara.

1

5. Though

That seemed unlikely, though.
Ayesha was wrong, though.
I hear quickly, though not as quickly as you.

6. However

The house, however, was the opposite.
Swords, however, are different.
Aima, however, was unaffected.

1

7. Therefore

~~I~~ am sleepy, therefore, I am going to bed.
I apologized, therefor.
I read a book, therefore, I learned new words

2

8. Well

The boy fell into the well. ✗
I cannot breathe very well. ✓
The old laptop still work well. ✗

9. Oh

oh. no The park was close.
oh my God! I cannot believe.
oh! we lose.

10. Really

Urooj really annoyed me.
I felt really nervous before interview.
Really you are so gorgeous.

11. I think

I think this is yours.
I think it is not a safe place.
I think she is selfish.

12. Also

She also had a sense of responsibility about it.
He also initiated it.
She also got a Prize.

13. Then

I seen you at five, then.
Then she was falling again.
Then why are you crying?

14. Instead

Instead he said nothing.
Instead of calling him she texted him.
Instead she felt wonderfully content.

15. Perhaps

Perhaps the ghost will also return.
Perhaps he had an unfair advantage.
Perhaps the only wild friend we have.

EACHING DISCOURSE MARKERS AT HIGHER SECONDARY SCHOOL EVEL USING DATA DRIVEN LEARNING APPROACH: AN EXPERIMENTAL STUDY

Delayed Post Test

Essay

15 8

Name:

Role Of Education

Education play a very important role in our life. Education is a word that teaches a person to live life. However in this world many still do not get eduction. Islam teaches that education whould be equally for both men and women because they make a healthy and educated society. Education give knowledge without education we cannot succeed in our life.

Education build the confidence and still therefore it helps us build the career. It also helps to learn many new and good things, so it member is beter citizen as well as good humans.

'Seeking Knowledge is the duty of Every Muslim':-

Education perpars us to face the world. Education teaches us

to be polite. Education is the Property that cannot be stolen and increase more and more by reading. (instead) of decreasing.

Education bring happiness in our life. And educated Person is always respected everywhere.

"Education moves from darkness to lightness"

Education is useful in every area of our life. In our Society uneducated Person is not treated with such respect and Protect as compared to an educated Person.

Finally Education is such a thing which removes the animality of a man and teaches him to live a meaningful life. Education improves our skills, Personality and Knowledge.

Therefore, it empowers the humans and prepare them to face the Chalsanges of life efficiently As he know that it is said.

'Seek Knowledge and Spread' :-