교육학석사학위논문

The Effects of Computer-Assisted Data-Driven Learning on Distinguishing Synonyms: A Case of Korean EFL Middle School Students

한국 중학생의 영어 유의어 학습을 위한 컴퓨터 활용 언어자료기반학습(DDL) 효과 연구

2023년 8월

서울대학교 대학원

외국어교육과 영어전공

정 하 영

The Effects of Computer-Assisted Data-Driven Learning on Distinguishing Synonyms: A Case of Korean EFL Middle School Students

by

Hayoung Cheong

A Thesis Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in Education

At the
Graduate School of Seoul National University

August 2023

The Effects of Computer-Assisted Data-Driven Learning on Distinguishing Synonyms: A Case of Korean EFL Middle School Students

지도교수 오 선 영

이 논문을 교육학 석사 학위논문으로 제출함

2023년 8월

서울대학교 대학원

외국어교육과 영어전공

정 하 영

정하영의 석사학위논문을 인준함

2023년 8월

위 원 장 _____

부위원장 _____

위 원 _____

The Effects of Computer-Assisted Data-Driven Learning on Distinguishing Synonyms: A Case of Korean EFL Middle School Students

APPROVED BY THESIS COMMITTEE:

_____

KITAEK KIM

_____

IN CHULL JANG

_____

SUN-YOUNG OH

ABSTRACT

The Effects of Computer-Assisted Data-Driven Learning on Distinguishing Synonyms: A Case of Korean EFL Middle School Students

Hayoung Cheong

English Major, Dept. of Foreign Language Education

Graduate School of Seoul National University

The comprehension and acquisition of vocabulary play a fundamental role in the development of communicative competence when acquiring the English language. While common words pose fewer difficulties, second language (L2) learners often struggle with distinguishing synonyms, which are words that share the same meaning in their first language (L1) but exhibit different grammatical or collocational constraints. The present study attempts to implement Data-Driven Learning (DDL), utilizing corpus data, to facilitate the understanding of distinctions between pairs of synonyms. Previous studies on DDL primarily focused on adult learners, as it is deemed more demanding for younger learners. Additionally, due to factors such as learners' proficiency levels and limitations posed by technology and the learning environment, many prior DDL studies have employed edited corpus

data.

Therefore, the present study aims to investigate the effectiveness of computer-assisted DDL on distinguishing synonymous verbs in case of Korean EFL middle school students. Given that individual computer devices were provided to 1st grade middle school students in Seoul starting from 2022, the participants in this study were asked to search authentic usages of synonyms from raw corpus data and discover the differences between the words themselves. Based on this process, the effectiveness of computer-assisted DDL along with learners' perception, attitude and changes were discussed.

Twenty-nine middle school 1st grade students in Seoul participated in this study. All participants took a pretest and had a training session prior to the experiment. Then, the students individually completed the computer-assisted DDL tasks for three weeks, focusing on distinguishing five pairs of synonyms. Following the completion of the experiment, the participants took a posttest and were asked to respond to a questionnaire. Additionally, individual interviews were conducted with three participants.

The results from the experiment showed the effectiveness of computer-assisted DDL on various aspects. First, the learners' ability to distinguish synonymous verbs developed. Second, the learners had positive perception toward computer-assisted DDL, particularly highlighting the intriguing nature of utilizing

computers during classroom activities. However, the individual interview analysis revealed variations among learners based on their proficiency levels. The high proficiency learner successfully completed the assigned tasks without encountering significant difficulties, whereas the intermediate learner required scaffolding and guidance from the instructor. The low-proficiency learners faced challenges in comprehending the computer-assisted DDL tasks, thus consistently relying on scaffolded support. These outcomes underscore the essential role of training sessions and scaffolding in computer-assisted DDL instruction for pre-tertiary learners.

Despite the small sample size and methodological limitation, the present study contributes to proving the effectiveness of computer-assisted DDL on distinguishing synonyms. The results of this study imply the possibility of implementing DDL into Korean middle school English lessons and utilizing corpus data as a learning tool.

Key Words: Corpus, Corpus-based Learning, Data-Driven Learning (DDL), Computer-Assisted Language Learning (CALL), Synonyms, Vocabulary Learning

Student Number: 2021-20625

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.
INTRODUCTION

This study aims to examine the effects of data-driven learning (DDL) on distinguishing synonyms in the case of Korean EFL middle school students, focusing on both the cognitive and affective domains. The present chapter introduces the background and purpose of the study, along with the necessity of conducting the current study. Section 1.1 presents the background and purpose of the study, followed by research questions in Section 1.2. Lastly, Section 1.3 outlines the organization of the thesis.

1.1. Background and Purpose of the Study

Despite the challenges faced by English speakers and learners, including native and non-native individuals (Friginal, 2018), distinguishing or differentiating synonyms is known to be a complex task. Particularly, EFL learners often encounter difficulties when discerning and employing synonyms (Liu & Zhong, 2016; Alanazi, 2017), as these words share similar translated meanings in both their L1 and L2 (Kim, 2020), but may differ in nuance and connotation (Inkpen & Hirst, 2006). Nevertheless, synonyms play a crucial role in effective and accurate communication, enabling speakers to convey shades of meaning (Edmonds & Hirst, 2002; Liu, 2010).

Numerous studies conducted within the EFL context have explored learners' usage of synonyms and revealed the challenges they face in using such words appropriately (Jung et al., 2006; Jung et al., 2007; Lee, 2011; Liu & Zhong, 2016; Alanazi, 2017). Building upon these findings, the primary objective of the present study is to discover an effective teaching method for distinguishing synonyms in Korean EFL middle school context.

In the context of Korean secondary schools, the relative importance of teaching grammatical knowledge compared to teaching vocabulary knowledge has been a subject of consideration (Jung et al., 2006). It has been observed that students often face the requirement of memorizing Korean definitions of individual English words without gaining a comprehensive understanding of their authentic usage, consequently leading to a limited ability to effectively employ appropriate vocabulary within contextual frameworks (Jung et al., 2006). This issue is particularly pronounced in Korean English textbooks, which fail to sufficiently provide authentic examples of English synonyms. As a result, there arises a pedagogical necessity to incorporate instruction on the accurate differentiation and appropriate usage of English synonyms.

Teaching how to use synonyms precisely differs from teaching vocabulary in a discrete sense, as providing definitions of certain synonyms from various dictionaries does not help learners to grasp the differences and nuances (Friginal,

2018). With the development of corpus linguistic, corpora have become an important and effective tool for understanding and learning synonyms as many previous corpus-based studies of synonyms have shown (Gries & Otani, 2010; Liu, 2010; Liu & Espino, 2012; Liu & Zhong 2016). Empirical evidence has substantiated the efficacy of utilizing corpora as a means of enhancing English learners' understanding of synonyms (Yeh et al., 2007).

Indisputably, corpus-based activities constitute an advantageous methodology for the instruction of vocabulary in a broader context. In detail, corpus-based data-driven learning (DDL) is argued to outperform traditional methods in vocabulary instruction (Boulton, 2008; Chujo et al., 2012; Frankenberg-Garcia, 2014). DDL is an inductive approach, leading learners to discover facts and rules about the language themselves based on the authentic examples (Johns, 1997). When applying DDL in learning procedures, learners are repeatedly exposed to various contexts so that they can discover and learn the meaning and usage of the words with better retention (Lin & Lee, 2019). Despite the advantages of applying DDL into language classes and the increased popularity of DDL over the past 20 years (Papaioannou et al., 2020), however, the meta-analysis of DDL studies by Boulton and Cobb (2017) showed that only 10 out of 88 studies dealt with secondary school students. DDL was mostly applied to adult learners of language as DDL is argued to be suitable for intermediate and advanced learners only (Gliquin &

Granger, 2010). This tendency is also in line with Korean EFL education (Lee, 2011; Kang, 2019, Kim, 2020), where many previous studies proved the effectiveness of DDL focusing on adult learners of English.

In addition, most of the previous studies of DDL used paper-based, edited concordances as learning materials due to lack of IT facilities (Lee, 2013; Lee et al., 2019; Lin & Lee, 2019; Kim, 2020). The educational environment was not yet prepared for students to use their own computer during class so many teachers prepared paper-based materials for DDL including edited concordances. Although such edited concordances may benefit learners with limited proficiency or at a novice level by enhancing comprehension, these resources inherently offer restricted data and compromise the essential authenticity required by DDL. Consequently, the adoption of raw corpus data in conjunction with computer technology has been advocated as a necessary measure.

Recently, the educational condition of Korea has witnessed notable changes. Digital competence and autonomy are being emphasized throughout the 2022 revised national education curriculum (Ministry of Education, 2021) and middle school 1st grade students in Seoul are currently receiving individual computer device, such as tablet PC or chrome book (Seoul Metropolitan Office of Education, 2022), for educational purposes. These trends necessitate a reevaluation of the application of DDL in the classroom. The evolving educational environment enables

the provision of raw corpus data in data-driven learning, thereby empowering students to assume the role of a 'language detective', as proposed by Johns (1997). The utilization of raw corpus data in DDL activities aligns with computer-assisted language learning (CALL), which emerged in the 1990s with the advancements in Internet and computer-mediated communication (Lim & Aryadoust, 2021). CALL has since expanded to encompass various domains of language learning, including corpora and data-driven learning (Farr & Murray, 2016). Previous studies have explored the development of student autonomy (Smith & Craig, 2013; Mutlu & Eroz-Tuga, 2013) and have demonstrated positive attitudes towards learning (Lasagabaster & Sierra, 2003) resulting from the integration of CALL into instructional practices.

Accordingly, the present study aims to examine the effects of computer-assisted data-driven learning on Korean EFL middle school learners, specifically focusing on distinguishing five pairs of synonymous verbs: *speak/talk, say/tell, hear/listen, end/finish, look/see/watch*, the verbs chosen from the basic vocabulary list for middle school students in 2015 revised National Curriculum of English. Additionally, the study will discuss the feasibility of incorporating educational devices during classroom activities. Finally, this research investigates the perception and attitudes of Korean middle school students towards corpus-based data-driven learning facilitated by computers. Considering all the changes within the educational

environment, data-driven learning using raw corpus through computer is expected to be actively adopted in schools with positive effects.

1.2. Research Questions

The present study investigates the effects of computer-assisted DDL on distinguishing five pairs of verb synonyms. In addition, it examines Korean EFL middle school students' attitudes toward computer-assisted DDL and their changes in participation throughout the lessons. The research questions for the present study are as follows:

1. Do Korean EFL middle school students improve their comprehension ability to distinguish synonymous verbs through computer-assisted DDL?
2. How do learners perceive computer-assisted DDL?
3. What changes do learners experience in their cognitive and affective domains through computer-assisted DDL?

1.3. Organization of the Thesis

The current study consists of six chapters. As previously mentioned, Chapter 1 explains the background and purpose of the study, along with the three research questions that are going to be examined throughout the study. Chapter 2 reviews the literature about synonyms in English and L2 learning of synonyms. The

two main approaches of teaching and learning languages applied in this study, DDL and CALL, are also discussed.

Chapter 3 introduces the methodological approach of this study. First, profiles of the participants and the setting of the study are provided. Next, the target items selection and the instruments of the present study are stated including the pretest and posttest, learning materials, and survey. Additionally, the procedures of the specific lesson and interviews of the present study are explained. Finally, methods for data collection and analysis are presented, categorized by quantitative and qualitative approach.

Chapter 4 asserts the results of the current study. Key findings are discussed based on the analysis of both quantitative and qualitative research data. Chapter 5 discusses a detailed analysis and possible interpretations of the findings regarding the research questions stated in Chapter 1. Chapter 6 summarizes major findings and pedagogical implications along with limitations of the present study and finally suggests future research.

CHAPTER 2.
LITERATURE REVIEW

This chapter reviews the theoretical and conceptual background of the present study. Section 2.1 discusses the concept of synonyms in general and L2 learning of synonyms. In Section 2.2, the concept of data-driven learning is introduced along with its application in vocabulary teaching and learning in EFL contexts. Then, Section 2.3 expatiates an overview of computer-assisted language learning, followed by the spread of data-driven learning in computer-assisted language learning researches in Section 2.4. Lastly, Section 2.5 summarizes the chapter by presenting the research gap that this study intends to fill in.

2.1. Conceptualization of Synonyms

Synonyms can be categorized into absolute synonyms and near-synonyms. Absolute synonyms are substitutable in any context without changing the truth value or meaning of the word (Liu & Zhong, 2016). Near-synonyms are "not fully intersubstitutable, but vary in their shades of denotation or connotation, or in the components of meaning they emphasize" (Inkpen & Hirst, 2006, p.223). Despite the categorization of synonyms, many linguists (Stubbs, 2001; Edmonds & Hirst, 2002; Taylor, 2003; Moon, 2010; Liu & Zhong, 2016) argued that absolute synonyms are

rare and almost impossible to define. This has led to the understanding that "all synonyms are truly near-synonyms" (Liu & Zhong, 2016, p.260). In this sense, synonymy is a pervasive and important but difficult linguistic feature for language learners to fully acquire and use (Liu & Zhong, 2016; Kim, 2020).

The theoretical background of synonymy started from the lexical-semantic theory suggested by Firth (1957), Halliday (1966), and Sinclair (1966) that the meaning of a particular lexical item is largely influenced by the contextual features including collocates. Based on this theory, lexical synonymy is commonly understood as semantic similarity, without changing the perceived meaning of the context (Arppe & Järvikivi, 2007). Cruse (2000, p.156-160) also defined synonyms as the words "1) whose semantic similarities are more salient than their differences, 2) that do not primarily contrast with each other, 3) whose permissible differences must in general be either minor, background, or both". To analyze the characteristics of lexical synonymy, many studies of lexical synonymy have developed based on a corpus-based approach (Liu, 2013), which enables the detailed analysis of semantic differences among synonyms.

2.1.1. L2 Learning of Synonyms

Comprehending and acquiring words is one of the core factors of communicative competence as the intention of the speaker or writer is reflected in

word choices. Individuals should understand both the word's core meanings and the underlined meanings (Hunston, 2002) to understand the intention, attitudes, and beliefs of a speaker or writer. Unlike common words, the lexical choice becomes difficult when it relates to synonyms (Kim, 2020), not only for L2 learners but also for native speakers (Martin, 1984; Edmonds & Hirst, 2002; Liu & Zhong, 2016; Friginal, 2018) as those words share a core meaning but have different grammatical or collocational constraints.

Indeed, the challenge of differentiating synonyms has been extensively acknowledged among researchers investigating EFL contexts (Jung et al., 2006; Jung et al., 2007; Jung, 2009; Morley & Partington, 2009; Park, 2011; Wongkhan & Thienthong, 2021; Yevchuk, 2022). Previous studies have consistently indicated that EFL learners often encounter difficulties when selecting the appropriate synonyms within given contexts (Jung, 2009; Morley & Partington, 2009; Yevchuck, 2022) and make frequent errors when using synonyms (Jung et al., 2007). Park (2011), for instance, interviewed twenty-three Korean EFL undergraduates and discovered that most of the students had insufficient understanding of synonymy as they have not learned about the detailed usage and shades of meaning in synonyms. Many students depended on using thesauruses when identifying synonyms, which normally do not provide information about connotation or usage.

Moreover, the influence of English proficiency level and academic

experience on the challenges encountered by EFL learners has been acknowledged, as higher proficiency individuals have demonstrated superior performance in collocation and synonymy assessments compared to their lower proficiency counterparts (Wongkhan & Thienthong, 2021). Wongkhan and Thienthong (2021) specifically found that Thai undergraduate students with limited educational experience often resorted to guesswork when selecting words, struggling to provide substantiated justifications, whereas students with an extended duration of education tended to articulate their rationale for choosing specific words. However, another study showed that compared to L1 English speakers, even advanced EFL learners had insufficient knowledge of synonyms (Yevchuk, 2022).

Despite the difficulties of discriminating synonyms, many studies have agreed that synonyms are important for effective and accurate communication that includes expressing shades of meaning (Hatch & Brown, 1995; Edmonds & Hirst, 2002; Liu, 2010), emphasizing the necessity of paying special attention to synonymy in language learning contexts (Jung et al., 2007; Liu & Zhong, 2016; Yevchuk, 2022). In this sense, investigating effective language learning activities dealing with synonyms deserves attention, especially given that former teaching methods have been shown to be ineffective in this area. It was found, for instance, that the use of bilingual dictionaries, which were traditionally used in vocabulary learning, is not as helpful as previously thought since they normally do not provide

sufficient information about word usage, only emphasizing denotation (Partington, 1998; Xiao & McEnery, 2006). Overall, advancements in corpus linguistics have prompted changes in the teaching and learning activities of synonymy. Corpora have become an effective tool for understanding and learning synonyms, supported by many corpus-based studies of synonyms (Gries & Otani, 2010; Liu, 2010, 2013; Liu & Espino, 2012; Liu & Zhong 2016).

Although synonymy has been studied based on corpora, the experimental studies which apply corpora as an educational tool for learning synonyms are insufficient. Yeh et al. (2007) investigated the effectiveness of using corpora in enhancing EFL learners' knowledge of the selected five sets of synonyms including *important, beautiful, big, hard,* and *deep*. The result showed that using corpora enhanced learning of synonyms as corpora effectively supported the learners' understanding of the targeted collocational patterns. Recently, Kim (2020) conducted a qualitative analysis with six undergraduate Korean EFL learners. Using the Corpus of Contemporary American English (COCA), the participants were encouraged to identify the differences between synonyms – *demand/request, mend/repair, outcome/consequence, happen/take place*. Despite the differences in problem-solving style among the six participants, they all successfully identified the differences between synonyms using the corpus provided by the instructor.

2.2. Data-Driven Learning (DDL)

Data-driven learning (DDL), a pedagogical approach suggested by Tim Johns (1986, 1988, 1991) was developed with language learners' uses of corpus data in the 1980s. It is broadly defined as "an approach in which learners taken an inductive approach to examples of language" (Hunston, 2022, p.174). To be specific, DDL consists of applying "tools and techniques of corpus linguistics for pedagogical purposes" (Gliquin & Granger, 2010, p.359). Although DDL did not receive public attention at first, as computers and concordancers become increasingly available from the 1980s, DDL began to be applied widely in language education (Chambers, 2010; O'Keeffe, 2021). Individuals are required to access various corpora and concordancing software via web and online corpus tools (Hunston, 2022) which has supported the popularity of DDL over the past 20 years (Papaioannou et al., 2020).

The essence of DDL is inductive learning (Chambers, 2010) in that students act as 'language detectives' (Johns, 1997), discovering facts and rules about the language themselves based on the authentic examples. Consequently, the teacher becomes the facilitator of learning (Johns, 1997) who provides guidance and support to students. To conclude, DDL reverses the traditional roles of teachers and students. The following sections deal with DDL in detail including the theoretical background, types, advantages and disadvantages, and previous experimental DDL researches

with their results.

2.2.1. Theoretical Background of DDL

Although DDL studies were widely conducted over the last three decades, the focus was on measuring net learning through organizing the pretest and posttest instead of scrutinizing the nature of learning (O'Keeffe, 2021). In addition, Pérez-Paredes (2022) pointed out that many previous DDL studies placed emphasis on empirical research examining effectiveness and language gains, without considering their relation with a theoretical basis. He also argued that these tendencies prevent many language teachers from fully understanding DDL in the context of second language acquisition (SLA) and language education. As many researchers have argued for the necessity of analyzing DDL employing a theoretical underpinning for broader pedagogical application (Römer, 2006; Tribble, 2008; Pérez-Paredes, 2010), this section introduces DDL in connection with sociocultural theory and second language acquisition theory.

2.2.1.1. DDL and Sociocultural theory (SCT)

There are two different views on interpreting the pedagogical stance of DDL: the Constructivism-focused paradigm and the SCT-focused paradigm (O'Keeffe, 2021). Cobb (1999) initially proposed that constructivist pedagogies

could provide theoretical support for utilizing corpora in language learning. O'Keeffe (2021) further supported this idea by explaining that "concepts such as induction, inference, hypothesizing, learner-centeredness, and discovery learning" (p. 261) can elucidate both constructivism and DDL. Within the constructivism-focused paradigm, learners independently engage in a discovery learning process. However, criticisms have been raised regarding constructivism itself (McGroarty, 1998; Kirschner et al., 2006), as some reports indicate that many learners find independent process-oriented learning demanding and challenging.

With concerns that a constructivist view of DDL that supports student-led learning might result in incorrect inferencing or lack of discovery during the learning process, SCT paradigm arose in the field of DDL. The SCT paradigm introduces the concept of 'scaffolding' as a means of supporting learners in mastering challenging skills during the acquisition process (Bruner, 1978). Several studies (Cobb & Boulton, 2015; Flowerdew, 2009, 2015) have advocated for the inclusion of scaffolding in DDL activities to reduce cognitive demands on learners. Flowerdew (2009) particularly emphasizes the stage of intervention where instructors provide scaffolding during corpus-based activities. Additionally, the SCT paradigm values learner self-regulation and agency (O'Keeffe et al., 2007; Cobb & Boulton, 2015; Flowerdew, 2015), recognizing that learners should assume control of the entire learning process to become self-regulated rather than passive participants. Moreover,

mediation, a crucial concept in the SCT theory, posits that cognitive processes in individual learners are facilitated through language as a tool, which can be mediated by oneself, a teacher, or a peer (Swain, 2006). SCT emphasizes that mediation, whether in the form of dialogue or inner speech, leads to knowledge acquisition. In summary, while constructivism focuses on independent knowledge discovery, the SCT paradigm places emphasis on mediation or scaffolding by peers or a teacher to foster self-regulated learning (O'Keeffe, 2021).

Based on the theoretical concepts of SCT, O'Keeffe (2021) concretized the actual language class as follows. In this approach, the teacher takes on a more prominent role by selecting the target language item in advance and delivering pre-instruction on that specific language form using curated data and a designed corpus task. Unlike the constructivist-focused DDL, where students have complete autonomy and no predetermined target form chosen by the teacher, the SCT view of DDL involves a more teacher-controlled process that still fosters learners' self-regulation through mediation and scaffolding. Despite the potential for SCT to offer valuable insights into DDL research, its exploration in this context has not been extensively investigated (O'Keeffe, 2021; Pérez-Paredes, 2022). Therefore, there is a need for detailed analysis of DDL that compares experimental studies from an instructional perspective and addresses the key concerns of instructed SLA based on the pedagogical foundations of DDL.

2.2.1.2. DDL and Second Language Acquisition (SLA) theory

Even though DDL is widely applied in the field of second language education, there are only a few studies that have argued for the importance of integrating different second language acquisition (SLA) theories with DDL (Johansson, 2009; Flowerdew, 2015; O'Keeffe, 2021; Pérez-Paredes, 2022). O'Keeffe (2021) suggested that with a sufficient number of valuable DDL experimental research studies or meta-analyses, DDL can provide comprehensive explanations of the cognitive processes involved in SLA and establish connections between implicit and explicit learning processes.

Among the array of second language acquisition (SLA) theories, the 'Noticing Hypothesis' formulated by Schmidt has garnered frequent attention in the context of DDL studies (Chambers, 2010; Lee et al., 2019). The Noticing Hypothesis elucidates the significance of attention in L2 learning (Schmidt, 1990), asserting that learners must consciously attend to target items in the input. In contrast to emphasizing implicit learning in L2 acquisition, Schmidt emphasizes the necessity of conscious awareness of linguistic forms. In line with the fundamental tenets of the Noticing Hypothesis, the concept of 'attention' is comprehended in a broader sense (O'Keeffe, 2021). In fact, several preceding studies have established a correlation between noticing and attention, as DDL experimental groups demonstrated heightened levels of noticing resulting from conscious or

subconscious attention (Boulton, 2010; Shi, 2014). DDL studies exhibit a close association with the Noticing Hypothesis, as they employ diverse corpora examples to facilitate learners' awareness of linguistic patterns (Chambers, 2010). Moreover, the strong interface position of instructed SLA which posits that explicit knowledge can be internalized and transform into learners' implicit knowledge (Han & Finneran, 2013), offers a compelling explanation in the context of DDL and the noticing hypothesis. Explicitly attending to the target form ultimately paves the way for implicit learning processes (O'Keeffe, 2021).

The learners' attention and noticing depend on how frequent and salient the input is, which illustrates another aspect of SLA theory, the importance of 'frequency' (Chambers, 2010). Frequency refers to "the number of times a learner has to encounter an aspect of language use to be aware of it and to be able to use it" (Chambers, 2010, p. 354) in the language-learning context. Learning is largely influenced and determined by the frequency of exposure to constructions (Ellis, 2006; O'Keeffe, 2021). In DDL, learners access corpora (concordances) and give attention to frequent patterns of examples. The learners' exposure to frequent repetition of the target item has been proven to be related to successful acquisition of lexical knowledge (Indrarathne et al., 2018), as learners become able to integrate the frequent examples of actual usages into their own language use.

Lastly, the concept of 'Input Enhancement' (Sharwood Smith, 1981) and

‘Involvement Load Hypothesis’ (Laufer & Hulstijin, 2001) may be linked to DDL. According to Sharwood Smith (1991), input enhancement is defined as “the process by which language input becomes salient to learners” (p.118). Input Enhancement is realized with Key-Word-In-Context (KWIC) in DDL, as keywords of concordances are obvious and clear enough to bring learners’ conscious attention (Wicher, 2020). Based on the characteristic of DDL that shows higher involvement in concordances, Lee et al. (2019) expanded this characteristic into Involvement Load Hypothesis. The fundamental argument of this hypothesis is that when there is more involvement loaded during the learning process of lexical items, learners will retain the word more easily. To conclude, O’Keeffe (2021) suggested that as learners are repeatedly engaged with certain input (input enhancement) in DDL, learners might easily learn and remember the lexical item (involvement load hypothesis). However, the same study pointed out a dearth of studies related to these hypotheses conducted by DDL researchers, with the suggestion of exploring this field more extensively with language aspects other than vocabulary.

2.2.2. Conceptualization of DDL

As DDL was mainly used in language-oriented research at first (Pérez-Paredes, 2010), the usage of corpora and DDL did not become mainstream in the foreign language education field until 2011 (Huang, 2011). However, the focus of

DDL switched from corpus linguistics to language pedagogy from 2011 (Boulton & Pérez-Paredes, 2014) and resulted in an increasing number of empirical DDL studies (Boulton & Cobb, 2017). In fact, Pérez-Paredes (2022, p.38) claimed that "DDL is trying to meet the needs of an ever-increasing number of learning contexts".

Although DDL implies the use of tools such as computers and computer software to access concordancer, indirect access of corpus data edited and provided by the instructor is considered as a different type of DDL. Gabrielatos (2005) categorized DDL into hard and soft versions. The hard version of DDL is marked by direct access to computer facilities and raw corpus data. This version of DDL embraces discovery-based learning, achieving the fundamental purpose of DDL as suggested by Johns (1991), which ultimately maximizes students' motivation of learning language (Huston, 2002). However, this hard version of DDL can be burdensome to learners. As there may be an overwhelming amount of corpus data (Varley, 2009), novice or young learners may not acquire what the instructor intended throughout the class (Huston, 2002). In contrast, the soft version of DDL refers to employing corpus-based materials selected and edited by the instructor. The soft version of DDL is closer to the teacher-led end where the instructor has more control and as a result, learners experience reduced burdens (Gabrielatos, 2005). Low proficiency or novice learners, in particular, may be better able to understand the class materials and learn the target items more effectively (Huston,

2002). However, the soft version of DDL offers limited data and, in this respect, learners may become potentially less motivated. Within the soft version, DDL becomes one of the tangential activities (Huston, 2002) Indeed, Gabrielatos (2005) pointed out the fact that the soft version with edited concordance lines should be provided to learners with the understanding that these do not represent a perfect frequency of a language item. Boulton (2009) also supported this idea by expressing dislike toward the soft version of DDL, as it undermines one of the primary advantages of DDL, authenticity.

The choice between the hard or soft version of DDL depends on two factors (Gliquin & Granger, 2010). First, the availability of the necessary hardware and software for applying the hard version of DDL is an important factor. The problems of facilities and logistics are one of the biggest limitations of DDL and will be discussed later. Second, the learners' level is another important factor. Although DDL is possible with all learners (Boulton & Cobb, 2017), a specific methodology needs to be adapted based on the learner's level (Hadley, 2002). Many researchers have found that typical DDL can be difficult for less proficient learners (Hunston, 2002; Gliquin & Granger, 2010; Cobb & Boulton, 2015; Wicher, 2020). This led to the suggestion by Charles (2007) that beginners or less proficient learners should work on a computer during class and study themselves at home with edited and selected materials on paper. Boulton (2008) also recommended applying the soft

version of DDL to less proficient learners. As each presentation has its obvious advantages and limitations (Gliquin & Granger, 2010), there have been many arguments about applying which version of DDL to the learners. Recently, Boulton and Cobb (2017) noted that there is an evolution toward the hard version of practices in DDL over the years. However, Meunier (2019) stated that the current tendency of DDL has not taken an actual digital turn and recommends further experimental studies based on the hard version of DDL with integrating new tools and tasks.

Regardless of the different versions of DDL, the advantages and limitations of DDL have been widely discussed. To begin with, DDL presents several advantages. First, DDL provides authenticity to language learners (Johns, 1997; Chambers, 2010; Gliquin & Granger, 2010). By accessing corpora, learners are exposed to a large number of actual language uses including contextual clues (Frankenberg-Garcia, 2014) that cannot be found in ordinary textbooks or dictionaries (Chambers, 2010). Gliquin and Granger (2010, p.359) pointed out that the exposure to authentic examples leads "to vocabulary expansion or heightened awareness of language patterns".

Second, DDL encourages learners to become active and autonomous (Johns, 1986, 1997; Chambers, 2010; Gliquin & Granger, 2010; Szudarski, 2018; Liontou, 2020). An element of discovery, which is implied in DDL, makes learning more interesting and motivating (Gliquin & Granger, 2010), ultimately leading to more

involvement in the learning process. In fact, previous studies described learners as travellers (Bernardini, 2001), researchers or detectives (Johns, 1997), which all emphasized the role of active learner and participant. Moreover, the characteristic of learner-led (Gabrielatos, 2005) discovery learning encourages learners to not only have more freedom but also have more responsibility for their own learning. By experiencing inductive learning through DDL, the learners take a more autonomous and reflective role. Thus, they become empowered (Mair, 2002) and self-regulated (O'Keeffe et al., 2007), having high confidence, self-esteem, and agency in their learning process. This second advantage is linked to the change in the traditional roles of teachers and students (Kim, 2020). Originally, teachers were considered as delivers of knowledge and learners as receivers. However, when applying DDL, the teachers become facilitators of the learning process (Chambers, 2010). The teachers define and provide meaningful context to the students so that the students can acquire knowledge by involving in DDL activities and become researchers. In this aspect, DDL suggests a new form of student-led language learning (Kim, 2020).

Lastly, DDL helps learners to acquire many learning skills which can be transferred to other fields of study (Gliquin & Granger, 2010). O'Sullivan (2007) suggested seventeen different general cognitive skills that learners can acquire through DDL, including 'observing, noticing, reasoning, analyzing, interpreting, reflecting, making inferences, differentiating, and verifying' (p. 277). Also,

according to Boulton and Cobb (2017), DDL results in the development of language sensitivity, noticing, and induction. These skills go beyond acquiring knowledge of certain language functions, such as vocabulary and grammar. According to Boulton (2009), DDL especially shows long-term effects on the development of general skills.

Despite its numerous benefits in language learning, DDL is not without limitations, as highlighted by several studies. First and foremost, one of the major challenges of DDL pertains to logistics (Gliquin & Granger, 2010; Pérez-Paredes, 2020; Schaeffer-Lacroix, 2020). Chambers and Bax (2006) define logistics as the availability of equipment, resources, and suitable classroom settings, particularly in the context of employing the more resource-intensive approach to DDL. Successful implementation of DDL necessitates technological equipment, such as computers, and software featuring corpus data. However, the acquisition of such hardware and software can incur substantial costs for schools and educational institutions (Gliquin & Granger, 2010), thus emerging as a primary constraint to conducting DDL in educational settings. While some freely available corpora and software options do exist, they often come with limitations and may not cater to learners with varying proficiency levels. The problem of logistics minimizes when applying the soft version of DDL, but as previously mentioned, the soft version of DDL does not guarantee the greatest advantage, authenticity to learners. Additionally, the soft

version of DDL causes the problem of time, as teachers have to create and edit the corpus data themselves. This time-consuming nature works as an obstacle to the implementation of DDL (Gliquin & Granger, 2010; Vyatkina & Boulton, 2017).

Second, the lack of teachers' understanding of DDL including corpus tools is another frequently mentioned problem (Gliquin & Granger, 2010; Cobb & Boulton, 2015; Boulton & Cobb, 2017; Crosthwaite, 2020). Many teachers do not have sufficient information about a new learning environment of DDL, such as which corpora are available or how to use corpora in the classroom due to the lack of training for them (McCarthy, 2008; Crosthwaite, 2020). In fact, according to a recent survey, many language teachers in Spain and UK answered that they were not familiar with corpora when teaching languages (Pérez-Paredes et al., 2018). Although DDL is a student-centered approach, the teachers should first be an expert to corpora before introducing them to learners (Mauranen, 2004). Therefore, many researchers (Mukherjee, 2006; Braun, 2007; Breyer, 2009; Chambers et al., 2011; Leńko-Szymańska, 2017) emphasized the necessity of combining DDL into teacher education programs. In another case, some teachers do not prefer to adopt DDL methodology due to the skepticism toward efficiency of DDL (Gliquin & Granger, 2010). For this tendency, Crosthwaite (2020) asserted a lack of language teacher's constructivist beliefs. He pointed out that it is important to persuade teachers that DDL not only has the potential of bringing benefits to the students but also supports

the professional development of teachers themselves. As DDL requires little control of teachers during the lesson compared to traditional teaching methodology, teachers should change their point of view and let students learn autonomously (Boulton, 2009; Gliquin & Granger, 2010).

Lastly, as students are not familiar with DDL (Boulton & Cobb, 2017; Crosthwaite, 2020), they tend to show negative reactions (Chambers, 2010; Gliquin & Granger, 2010; Hirata, 2020). As there are overwhelming amounts of corpus data, some learners fail to find any patterns or rules (Whistle, 1999; Varley, 2009). Moreover, some previous studies noted applying DDL and analyzing corpus data as time-consuming, discouraging, laborious, and tedious (Yoon & Hirvela, 2004; Chambers, 2005; Boulton, 2009) despite some helpful and confidence-boosting aspects (Kennedy & Miceli, 2001). To overcome the difficulties and change students' points of view, learner training as well as learner-friendly corpus tool should be developed (Forti, 2008; Crosthwaite, 2020). Also, further support to students is required (Chang, 2012) so that pre-tertiary learners can successfully analyze corpus data during DDL.

Despite the limitations mentioned above, it is clear that DDL is a promising methodology of teaching. DDL facilitates learners' exposure to authentic language, enhances motivation, and fosters the development of skills beyond the linguistic knowledge (Chambers, 2010; Gliquin & Granger, 2010). Moreover, DDL aims to

foster the independent acquisition of linguistic knowledge through an inductive learning process (Kim, 2020; O'Keeffe, 2021). Building upon these aspects, Mizumoto and Chujo (2015) assert the superiority of DDL over other learning approaches. However, Gabrielatos (2005) emphasized the dangers of corpus worship, where a corpus is considered as successful teaching method in any context. Gliquin and Granger (2010) supported this idea by mentioning that DDL approach may not be an effective teaching methodology for all aspects of language. In their future study, they expanded this idea into the term 'DDL worship' (Gliquin & Granger, 2022). Thus, it is crucial to acknowledge learners' diverse learning styles and preferences when implementing DDL, as "a key word in DDL is variety" (Gliquin & Granger, 2010, p. 365).

2.2.3. Implementation of DDL in Language Learning Contexts

DDL has been used in various language learning contexts worldwide. Discovery learning approach using corpus data has been reported to show generally positive effects both in grammar (Hong & Oh, 2008; Liu & Jiang, 2009; Saeedakhtar et al., 2020) and vocabulary acquisition (Frankenberg-Garcia, 2014; Kim, 2020; Lee & Lee, 2010; Lee & Lin, 2019; Lee et al., 2019; Li, 2017; Pérez-Paredes, 2019) regardless of L1 and L2. The meta-analysis of DDL by Boulton and Cobb (2017) concluded the previous studies by noting that "DDL works pretty well in almost any

context where it has been extensively tried" (p. 386). But they added that DDL is most appropriate in foreign language vocabulary contexts for students with intermediate or advanced levels with hands-on concordancing compared to paper-based material. In addition, Lee et al. (2019) analyzed the effects of corpus use on L2 vocabulary learning and reported that DDL showed a medium-size effect with enhancing in-depth knowledge for the learners having at least an intermediate level of L2 proficiency. The interesting point of a meta-analysis by Lee et al. (2019) was that DDL can be effective without prior training.

DDL affords not only effectiveness but also guarantees language learners' positive attitudes (Crosthwaite & Stell, 2020; Pérez-Paredes, 2022; Szudarski, 2020). Recently, Pérez-Paredes (2022) found that 69% of DDL journal articles between 2011-2015 asked learners to express their attitudes towards using corpus mainly through questionnaires and some interviews. Using these replies, he discovered that learners valued DDL for its "usefulness" (p. 46), mostly for learning vocabulary and collocational behavior, followed by writing, register awareness, and speaking.

Despite these positive effects of DDL in the language learning context, applying DDL and corpora is still limited in several aspects. First and foremost, the vast majority of DDL studies have targeted tertiary learners, the limitation pointed out by many previous researchers (Boulton, 2008, 2009, 2010; Boulton & Cobb, 2017; Crosthwaite, 2020; Hirata, 2020; Pérez-Paredes, 2020, 2022; Tyne, 2012;

Wicher, 2020). According to Boulton and Cobb (2017), only ten out of eighty-eight studies explored the use of DDL among pre-tertiary learners and the meta-analysis of Pérez-Paredes (2022) concluded that only two papers out of thirty-two focused on young learners. Lee et al. (2019) did not even include age variables in their meta-analysis. As DDL studies tended to be limited only to adult learners with advanced language proficiency (Wicher, 2020), researchers have tried to explain this tendency for various reasons. For instance, there may exist some apprehension and fear on the part of learners and teachers (Boulton, 2009; Hirata, 2020) as learners are unfamiliar with the process of self-guided learning in DDL and the usage of corpus tools, while teachers normally consider that their young students would not be able to handle DDL so that they cannot successfully achieve their goals through DDL. Also, young learners' insufficient level of language proficiency (Boulton, 2010) and inappropriate corpus resources and tools for young learners to use (Hirata, 2020; Pérez-Paredes, 2020; Varley, 2009) prevent the development of DDL studies in secondary schools.

Given the limitation that only a small number of secondary language teachers have implemented DDL in their classrooms, there has been a growing body of research focusing on DDL for pre-tertiary learners (Boontam & Phoocharoensil, 2018; Crosthwaite & Steeples, 2022; Moon & Oh, 2018; Papaioannou et al., 2020; Pérez-Paredes, 2020; Saeedakhtar et al., 2020; Szudarski, 2020; Vyatkina, 2016).

These studies have provided evidence that DDL is effective even for learners at the beginning level of language proficiency, supporting Boulton's (2010) assertion that corpus analysis can be successfully utilized with lower-proficiency learners. However, there are still certain constraints associated with DDL for pre-tertiary learners, particularly when they are exposed to pre-prepared and edited corpus data, which can pose challenges considering their language proficiency. This aspect will be further discussed in Section 2.4, which addresses both DDL and CALL.

When applying DDL to pre-tertiary learners, it is important to provide appropriate DDL training because the majority of students are unfamiliar with DDL methodology and corpus data (Papaioannou et al., 2020; Liontou, 2020). According to Liontou (2020), initial training sessions should focus on teaching pre-tertiary students how to search online corpora and analyze concordance lines. Over time, learners will become more autonomous and independent as they adapt to the DDL approach and make progress. Despite the increasing number of studies demonstrating the positive effects of corpus use in secondary school language learning contexts (Schaeffer-Lacroix, 2020), additional attention is recommended, such as the development of child-friendly corpus resources and raising awareness among secondary school teachers regarding corpus tools (Braun, 2007; Crosthwaite, 2020; Tyne, 2012), in order for DDL to become mainstream in secondary language education.

The second point is that the body of DDL research under analysis tends to favor quantitative research methods (Pérez-Paredes, 2022). According to Kim (2020), previous studies examining the effects of DDL or corpus-based learning are designed as involving large samples to provide evidence of the general effectiveness. However, for a better understanding of how learners benefit from DDL, qualitative studies should be conducted, at least in the form of semi-experimental or mixed-methods research (Pérez-Paredes, 2022). DDL studies require exploratory approaches or action research to increase the visibility of DDL across various languages, levels, and instructional contexts (Hanks, 2019). Recently, Kim (2020) conducted a qualitative study of DDL with six undergraduate participants in Korea. Based on in-depth interviews and recordings, Kim (2020) reported detailed behaviors between intermediate and advanced learners were different when discriminating synonyms using edited concordance lines. Intermediate students tended to focus on getting the correct answer without acknowledging the primary meaning of the words, while the advanced learners correctly differentiated the structural differences of the synonym pairs and even tested their previous knowledge based on the given data. Crosthwaite and Stell (2020) also conducted a qualitative DDL study where two Australian primary school students participated in a writing class. Based on the observational notes, screenshots, and interviews, this study successfully analyzed young students' reactions and behavior toward DDL, along

with their changes over the session. These studies support that qualitative study is necessary to examine how students complete the DDL task, how they change, and what they feel toward DDL in detail.

About 200 empirical DDL studies were conducted between 1989 and 2014 across various ranges of languages and contexts (Boulton & Cobb, 2017). DDL studies in Korea also reflected this tendency as many previous researches proved the positive aspects of DDL in the various fields of English. For further development, as previously mentioned, DDL studies should focus on pre-tertiary learners with hands-on concordances. Also, studies must include qualitative methodology. Chambers (2019) mentioned that there is still a research-practice gap in the field of DDL, as DDL is widely known to applied linguistics, but not to the teachers in school. For DDL to gain mainstream acceptance, this research-practice gap should be overcome resulting in the wide application of DDL into EFL contexts in secondary schools.

2.3. Computer-Assisted Language Learning (CALL)

Computer-Assisted Language Learning (CALL), "using and studying the applications of the computers in teaching and learning language" (Levy, 1997, p. 1), has been used in the language classroom since the late 1960s (Lim & Aryadoust, 2021). Since then, CALL has continuously progressed along with the development

of various technologies, such as computer software and applications (Levy, 1997). From the period of 1990's, with the introduction of the Internet, CALL has been reshaped and widely used for language learning and teaching (Lim & Aryadoust, 2021). CALL finally expanded to various areas, including corpora and data-driven learning (Farr & Murray, 2016). The following sections deal with the theoretical background of CALL, the advantages and disadvantages of CALL, and the implementation of CALL in language learning contexts.

2.3.1. Theoretical Background of CALL

In most cases, researchers are expected to use theoretical underpinnings to expand existing knowledge, build on, or refine theories in their research. Moreover, practitioners such as teachers and designers use theories to support their decisions about what to focus on or ignore in their actual teaching (Levy & Stockwell, 2006). However, it is a surprising fact that there are no prominent theories of CALL. Hubbard (2008) analyzed articles within the *CALICO Journal* database mentioning 'theory', a total of 166 articles for 25 years since 1983, and found that there is no dominant CALL theory and that the studies tended to show a wide range of theoretical underpinnings, including SLA theory, learning theory, sociocultural theory, and so on. Stockwell (2012) suggested that as CALL shows diversities and complexities, it is natural to have multiple theories.

Among various theories, similar to the theoretical underpinning of DDL, sociocultural theory is frequently described in relation to CALL (Chapelle, 2003; Warschauer, 2005). The concept of mediation suggested by Vygotsky (1981) explains the notion that tools or signs mediate all human activity. In other words, various tools, such as computers in CALL, fundamentally change human action. As new technologies become included in the learners' learning process, they work to transform the whole flow and structure of the learners' mental functions (Warschauer, 2005). Additionally, the concept of social learning, where learners develop and learn through interaction with others, including the language of others or responding to others' reactions (Warschauer, 2005), In CALL, learners incorporate certain linguistic chunks and refine their input from the authentic examples provided by new educational technologies (Warschauer, 2002). To conclude, based on sociocultural theory, the computer in CALL interacts with the learners as a tool to promote language learning and ultimately transform human behavior.

2.3.2. Conceptualization of CALL

In the early stages of CALL, Pederson (1987) summarized the research on education and CALL. One of the findings she concluded was that meaningful CALL practice is preferable, as students usually demonstrate positive attitudes toward

CALL. Moreover, the design of CALL tended to result in more learning compared to conventional teaching methodologies. As the interest in CALL expanded to language learning and teaching by providing abundant language learning resources, the advantages of CALL were stated by many researchers. Reinders and White (2010) synthesized these advantages and categorized them into organizational and pedagogical advantages. The first organizational advantage is 'access'. Learners can access CALL materials anytime and anywhere, even outside the classroom after school, so that they can revisit the content area (Choi, 1996). The second one is for the material developers. CALL materials are easy to change and update by using a computer. Also, developers and teachers can immediately share the materials with learners online. The last organizational advantage is related to the record-keeping and storage functions of computers. When conducting CALL, learners' progress in learning, including the test results, can be electronically stored so that teachers can retrieve them at any time. Fisher (2012) stated that the CALL environment is close to a controlled environment as students' actions are visible and they are less influenced by external forces compared to the typical classroom environment. This statement is supported by previous research (Glendenning & Howard, 2003; Pujolà, 2002) that found that using video recording software in CALL to capture and track students' actions provides a complete and objective record compared to teachers' direct observation. The advantage of storage and retrieval of learners' learning

behavior will be discussed further in the aspect of pedagogical advantages.

CALL has more pedagogical advantages than organizational advantages. First, CALL shapes a less stressful setting for students (Lasagabaster & Sierra, 2003) as it uses computers or new technologies during class. As CALL fosters a learner-friendly learning environment, students can actively participate during the lesson, which also encourages their motivation for learning (Kim, 2002; Felix, 2005a; Jang, 2012). Second, CALL materials facilitate interaction and language use (Reinders & White, 2010), which goes in line with sociocultural theory explained in Section 2.3.1. Various CALL programs encourage the interaction between the learner and the computer, which ultimately leads to abundant language use. The third and most outstanding benefit of CALL is related to the concept of 'learner empowerment'. As CALL enables recording, as previously mentioned in the organizational advantage of CALL, learners can monitor their behavior and progress. This ultimately develops learners' metacognitive awareness and helps them develop autonomy (Mutlu & Eroz-Tuga, 2013; Reinders & Darasawang, 2012; Smith & Craig, 2013). Having control over themselves, learners might become self-regulated, which goes in line with the advantage of DDL. To sum up, CALL offers the potential to empower learners as it guarantees free and easy access in learners' own time, provides greater control to learners (Reinders & Darasawang, 2012), and finally gives learners opportunities to work autonomously (Stockwell, 2012).

However, some limitations of CALL also exist. First, CALL cannot be successfully applied to the language learning classroom when technical and financial supports are limited (Shin & Son, 2007). Similar to the limitations of DDL, the logistics should be well-prepared prior to the actual implementation of CALL. Another possible limitation is related to digital literacy. For the successful implementation of CALL, both the instructor and the learner need digital literacy. Pedagogical and technical training for teachers and learners must be conducted in advance (Hubbard, 2004; Levy & Stockwell, 2006), including ways to use the tools and resources effectively. The one interesting point is the recent tendency among young learners. The use of various kinds of digital devices leads many younger students to possess 'digital wisdom' (Prensky, 2011), which may solve the problem of computer literacy skills. Lastly, the teachers' point of view might hinder the adaptation of CALL. Some teachers may reject using CALL in class not only because they lack CALL-related knowledge but also because of the irrelevant contents certain technology contains (Stockwell, 2012). Teachers should monitor the learners in CALL to ensure that they successfully achieve the learning goals. At the same time, they should admit learner varieties, as some learners will show unexpected behavior (Fischer, 2007; Tanaka-Ellis, 2010), such as doing things that they should not do or not doing things they should do.

2.3.3. Implementation of CALL

Davies et al. (2014) argued that the field of education has entered the era of Computer-Assisted Language Learning (CALL) due to the widespread utilization of digital tools for learning, not only in real-world settings but also in the context of language education. A recent survey conducted by Pérez-Paredes et al. (2018) illustrated that approximately 70% of higher and secondary language teachers in Spain and the UK incorporate digital tools, such as online platforms or web-based services, in their language classrooms. The advancement of technology has given rise to Digital Language Learning (DLL), which encompasses a diverse array of emerging digital technologies (Li & Lan, 2022). CALL serves as the foundation for DLL and has further expanded to include mobile-assisted language learning (MALL), virtual reality (VR), and digital game-based language learning (GBLL).

As CALL has become prevalent in language education, many previous studies examining the effectiveness of CALL have been conducted regardless of various language aspects in both L1 and L2 contexts (Allum, 2002; Felix, 2005b; Grgurović et al., 2013; Kulik, 2003; Liu et al., 2002; Nim Park & Son, 2009; Yi & Cha, 2016). For example, Allum (2002) conducted a comparative study with adult learners to justify using CALL is more effective than traditional teaching in some aspects. The result showed that CALL methodology is credible and sound so that it can be successfully applied to language classes.

Based on the results, meta-analyses of CALL have been evident throughout many years. Kulik (2003) organized the effectiveness of CALL in higher education, showing an overall positive instructional effect with meaningful effect sizes. In case of student performance, all of the CALL studies indicated moderate to large improvement. Grgurović et al. (2013) conducted an in-depth meta-analysis, analyzed 85 studies of CALL from 1970 to 2006. The overall analysis goes in line with previous studies but one distinct characteristic was found. There was an effect on proficiency level. Learners with advanced and intermediate levels did better in CALL-based lessons compared to beginners.

While previous studies on Computer-Assisted Language Learning (CALL) have demonstrated its positive effectiveness for learners, its widespread use has primarily been limited to higher education (Grgurović et al., 2013). To establish the generalizability of CALL's effectiveness across all age groups and proficiency levels, it is essential to conduct research on CALL in primary and secondary education. Taking into account the age factor, Yi and Cha (2016) conducted an experiment to examine the impact of CALL on the speaking proficiency of EFL learners. An interesting aspect of this research was the participation of 82 learners ranging in age from 4 (kindergarten group) to 50 (tertiary group). The findings indicated that CALL software was effective for all age groups, although young learners demonstrated more balanced and significant growth in speaking proficiency.

This study aligns with the concept of 'digital wisdom' (Prensky, 2011), which suggests that young learners can make informed decisions by utilizing and integrating technology into language learning (Liu et al., 2002). The study by Yi and Cha (2016) highlights the potential of applying CALL to learners in pre-tertiary education.

The successful implementation of Computer-Assisted Language Learning (CALL) for young learners necessitates training for both teachers and learners, which serves as a crucial factor. Several researchers have emphasized the importance of teacher training in integrating CALL effectively into language classes (Chamber & Bax, 2006; Pederson, 1987; Pérez-Paredes, 2022). Similarly, learner training for CALL has received increasing attention in the literature (Barrette, 2001; Hubbard, 2004, 2005; Hubbard & Romeo, 2012; Kolaitis et al., 2006; Levy & Stockwell, 2006; O'Bryan, 2008). This training process aims to enhance learners' technological competence, particularly for the purpose of second language acquisition. Barrette (2001) analyzed fourteen CALICO journals from 1997 and 1998, revealing limited evidence of learners' previous computer literacy and the provision of training on the applications. Therefore, Barrette (2001) emphasized the necessity of assessing learners' computer literacy and providing appropriate training for the effective use of CALL. In a subsequent review, Hubbard (2005) examined 78 research studies in CALL literature published between 2000 and 2004. The meta-

analysis showed that only 31% of the studies included basic learner training at the beginning, and 29% mentioned the need for additional training to achieve better results. Despite the perception that many young learners today possess technological competence due to their familiarity with various digital devices, they do not automatically possess the skills to effectively utilize digital tools and materials for language learning (Hubbard & Romeo, 2012). These previous studies collectively underscore the importance of pedagogical and technical learner training in CALL (Hubbard, 2004; Levy & Stockwell, 2006). Learner training for CALL is a valuable process (Hubbard & Romeo, 2012), resulting in positive outcomes such as improved performance (Hubbard, 2005; Kolaitis et al., 2006; Nim Park & Son, 2009).

Above all, it is important to keep in mind that technology itself does not facilitate language learning in CALL (Doughty & Long, 2003). Language learning depends on how technology is capitalized upon in various learning environments. In other words, the implementation of CALL into classroom is highly influenced by diverse factors, including both inside and outside of the classroom (Stockwell, 2012). Consequently, it is essential for researchers and teachers to acknowledge the existence of these diversities within the teaching environment to ensure successful research and practice in CALL.

2.4. Spread of DDL in CALL Researches

Although the definition of Digital Data-driven Learning, as discussed in Section 2.2, implies the use of computers or software such as concordancers, it is important to note that DDL studies do not always presuppose Computer-Assisted Language Learning (CALL). DDL is not yet considered a primary area of focus in CALL-related studies. For instance, in the meta-analysis of CALL conducted by Grgurović et al. (2013), references to DDL or language corpora were not found. This finding is further supported by a recent review on the prevalence of corpora and DDL in CALL research from 2011 to 2015 (Pérez-Paredes, 2022). Among the 759 published CALL papers reviewed, only 32 explored the use of DDL and corpora, accounting for a mere 4.2% of the total. These papers primarily focused on DDL-assisted writing, emphasizing learning and pedagogy rather than computer technology, often employing short-term experiments.

Pérez-Paredes (2022) also discovered that 94% of the studies were conducted at universities, where researchers had easier access to tertiary students. This observation suggests a tendency towards sample bias in DDL studies, as the use of hard versions of DDL is often deemed inappropriate for low-level learners (Boulton, 2009; Lee et al., 2017). Regarding training, 60% of the research papers highlighted the necessity and importance of training learners before the lesson to ensure effective performance during the experiment. However, technical issues such

as access to technology were not identified as impeding factors when applying hard versions of DDL. This implies that educational contexts designed to utilize computers and software as digital devices are becoming more prevalent.

Despite some positive responses towards DDL in CALL studies, the use of hard versions of DDL is still limited due to certain obstacles. Therefore, further research is needed to overcome these challenges and facilitate its wider adoption.

2.5. The Present Study

This chapter has provided a review of synonyms in English, as well as Digital Data-driven Learning (DDL) and Computer-Assisted Language Learning (CALL) respectively. Based on previous research, it is evident that the investigation of the hard version of DDL is still evolving, particularly in the case of pre-tertiary learners. Therefore, the current study aims to examine the effectiveness of computer-based DDL in distinguishing synonymous verbs for EFL middle school learners, with a focus on the following aspects.

The accurate discrimination of synonyms holds significance for EFL learners as it aids in enhancing lexical choice, including the understanding of connotations (Jung et al., 2007; Kim, 2020), ultimately leading to the development of English communicative competence and overall proficiency. However, research on the learning of L2 synonyms has been limited (Liu & Zhong, 2016), except for

corpus-based studies that analyze synonyms. Moreover, it is challenging to find studies utilizing corpora to teach synonyms to pre-tertiary learners (Yevchuk, 2022), as most corpus-based learning primarily involves university students.

Furthermore, previous DDL studies have predominantly focused on tertiary learners (Boulton & Cobb, 2017; Pérez-Paredes, 2022). The DDL studies conducted in Korea align with the meta-analysis of DDL studies by Boulton and Cobb (2017), which primarily focused on college students' use of corpora (Jung et al., 2007; Lee, 2011; Kang, 2019; Kim, 2020). However, an increasing number of studies have started to concentrate on data-driven learning for pre-tertiary learners recently (Boontam & Phoocharoensil, 2018; Crosthwaite & Steeples, 2022; Papaioannou et al., 2020; Pérez-Paredes, 2020). Hence, it is essential to investigate the effectiveness of DDL with a focus on Korean pre-tertiary learners to apply DDL in regular English classrooms.

Moreover, most previous studies have utilized paper-based, edited concordances as classroom materials for DDL due to the lack of facilities and learners' low proficiency (Kim, 2020; Lee, 2013; Lee et al., 2017; Lin & Lee, 2019). Since DDL is most effective in its hard version, which involves hands-on corpus usage (Boulton & Cobb, 2017), a study examining the effectiveness of the hard version of DDL (Saeedakhtar et al., 2020) should be conducted to encourage an active role as 'language detectives' (Johns, 1997) and foster learners' 'digital

wisdom' (Prensky, 2011).

Lastly, previous research examining the effects of corpus-based learning has primarily involved quantitative studies with large samples to provide evidence of the effectiveness of corpus-based data (Pérez-Paredes, 2022). However, the results from quantitative studies alone are insufficient to examine how EFL learners utilize corpus data, how they test their own hypotheses, and their actions during the process of corpus-based learning. To bridge this gap, the present study incorporates qualitative methods to investigate the trajectories of a relatively small number of learners using corpus data and explore the learners' 'nature of learning' (O'Keeffe, 2021, p. 262).

CHAPTER 3.
METHODOLOGY

This chapter illustrates the methodology and research design employed in the current study. A mixed-methods design employing both quantitative and qualitative approaches is adopted in the current study, as a number of previous DDL studies have tended to favor quantitative research methods (Pérez-Paredes, 2022). By adding a qualitative research element, this research leads to a greater understanding of underlying behaviors and examines the research questions proposed in Chapter 1. Section 3.1 provides the profiles and basic information about the participants of this present research, followed by the setting in Section 3.2. In Section 3.3, the target items of this study are explained. Then, Section 3.4 provides the instrument details including the pretest and posttest materials, learning materials, questionnaire are introduced. Then, Section 3.5 explains the procedures of the computer-assisted DDL instruction and the experiment. Finally, Section 3.6 discusses how the collected data are quantitatively and qualitatively analyzed.

3.1. Participants

This study involved twenty-nine EFL middle school 1st grade students who are attending J middle school located in Seoul, Korea. Prior to the study, the researcher distributed an explanation of the research, along with parental consent forms. Only those students who voluntarily agreed to participate and from whom written parental consent was received were included in the study. Out of a potential thirty-seven participants, six students disagreed to participate[1] and two students were absent for more than half of the experimental process, resulting in a final total of twenty-nine participants. Specifically, 16 of them were male, and 13 were female.

The participants were limited to middle school 1st grade students as only 1st graders received their own educational digital device (Seoul Metropolitan Office of Education, 2022), a Microsoft Surface Go 3. There were no other inclusion criteria, such as English proficiency, as the purpose of the present study was to examine the effectiveness of teaching methodology, computer-based DDL regarding all proficiency levels of EFL learners. All participants were native speakers of Korean and they were studying and learning English as their foreign language. They could read and understand English sentences in Middle School English 1 textbook[2].

A survey asking for demographic information (see Appendix 1) shows that

[1] Among six students who disagreed to participate, four students were student athletes.

[2] Kim, J.-W. (2018). *Middle School English 1*. Visang.

the participants started to learn English in the mean age of 7.72 (*SD* = 1.51), ranging from 5 to 10. Most of them (*n* = 26) have learned English in an instructional setting in Korea, with the exception of three students who studied abroad to learn English for less than 3 months. Only one student heard about the concept of corpus and no one had experienced using corpora as a language learning tool before this experiment. The participants attended 90 minutes long DDL classes per week for four weeks.

Among twenty-nine participants, three students volunteered to have an interview with the researcher after every session. Those students are Hanna, Irene, and Leo[3]. Hanna and Irene are female students, and Leo is a male student. Hanna started to learn English when she was seven years old, while Irene and Leo started to learn English when they were nine years old. All of them have studied English only in Korea.

3.2. Setting

As explained in Section 3.1, this study was conducted at J middle school located in Seoul, Korea. J middle school is a public middle school with a total of 724 students. Among them, there were 98 male and 121 female 1st grade students in November, 2022. One notable characteristic of J middle school is the existence of

[3] Pseudonyms are used to protect the participants' anonymity

an official middle school soccer club, which is somewhat uncommon for Korean public middle schools. Students enrolled in the club primarily focus on developing their soccer skills, often at the expense of other areas of study. Approximately 30 of the male students at J Middle School are student athletes. Given the school's emphasis on athletics and the time commitment required, it was expected that some parents and students might be hesitant to participate in any research that could be perceived as requiring additional effort.

All participants were enrolled in an English class as part of a 'Theme selection activity' during the 'Free Year Program', an educational policy in Korea for middle school 1st grade students. The Free Year Program was introduced as an extension of the Free Semester Program in 2018. This program encourages students to explore their own career paths without the burden of formal tests, such as midterms or final exams (Kim & Kim, 2021). As part of the Free Year Program, middle school 1st grade students have the opportunity to choose various activities, including arts, sports, club activities, and the theme selection activity (Ministry of Education, 2013), as elective courses. The experiment in this study was conducted within the framework of this elective course, particularly the course of learning English with various contents, allowing participants to engage in the English class without the pressures associated with grading students.

3.3. Target Items Selection

In the current study, the research focused on five pairs of synonymous verbs as the target items: *speak/talk, say/tell, hear/listen, end/finish, look/see/watch*. The selection of these verbs was influenced by the findings of Jung et al. (2006), which indicated that EFL learners encountered challenges in distinguishing synonymous verbs compared to synonymous adjectives. All of the selected verbs belong to the basic vocabulary list outlined in the 2015 revised National Curriculum of English, which is implemented for middle school 1st grade students in the year 2022. Considering the participants' grade level, it is recommended that these specific words be addressed during elementary school instruction. Each pair of verbs possesses the same meaning in the learners' first language (L1), but exhibits variations in terms of complement structure or usage, as depicted in Table 3.1.

First, although the verb *speak* and *talk* both have the same L1 (Korean) meaning, *speak* is used with language as in (1a) while *talk* is usually used with preposition and object as in (1b) and (1d). Moreover, the context of *speak* is generally one-way as in (1c) while *talk* refers to two-way communication as in (1d) (Dirven et al., 1982). The following examples from *Lextutor*, a concordancing tool utilized in this study, show the differences in detail.

Table 3.1.

Selected synonymous verbs and the usage

Synonym pairs		Usage	
1	Speak	1) used with language	2) one-way
	Talk	1) used with preposition (e.g., *to, about*)	2) two-way communication
2	Say	message-focused	
	Tell	1) addressee-focused	2) used with human being object
3	Hear	without attention	
	Listen	1) paying attention	2) used with preposition *to*
4	End	1) complete	2) used with human being subject
	Finish	1) incomplete	2) used with non-human subject
5	Look	1) paying attention	2) used with preposition *at*
	See	regarding physical sight	
	Watch	moving object (e.g., movie, TV)	

(1a) Most of them **speak** English. (used with language)

(1b) I need to **talk** to someone. (used with preposition)

(1c) He began to **speak** very quickly. (one-way communication)

(1d) Let's **talk** about other things. (two-way communication)

Second, the verb *say* has greater message focus while the verb *tell* not only focus on message but also focus on addressee (Dirven et al., 1982) despite their same meaning in L1. In specific, the verb *tell* is mostly followed by human being object. Example (2a) ~ (2d) from *Lextutor* shows the differences in detail.

(2a) Helen's mother **say**s that is rude. (message-focused)

(2b) Paul **say**s that he enjoyed this time in India. (message-focused)

(2c) Why didn't Nick **tell** me about you? (addressee-focused)

(2d) I **told** him that I was busy. (human-being object)

Third, the main difference between *hear* and *listen* is the 'attention' (Swan, 2005). We tend to *hear* something without attention as in (3a), but we *listen* to something with paying attention as in (3b). Also, normally *listen* is followed by preposition *to* when having an object as in (3c).

(3a) I began to **hear** a loud sound. (without attention)

(3b) Now, **listen** carefully, because this is very important. (paying attention)

(3c) I don't want to **listen** to her. (preposition *to*)

Fourth, according to Freed (1979), Sim (2010) and Huiying (2013), *finish* is used when something is being over and complete, usually with a human being subject. In contrast, *end* is used although something is incomplete, usually with a non-human subject. Example (4a) ~ (4d) shows the clear differences.

(4a) Stay and **finish** your dinner. (complete)

(4b) This game has not **end**ed yet. (incomplete)

(4c) He **finish**ed eating and got up from his table. (human being subject)

(4d) The story **end**ed happily for everybody. (non-human subject)

The last pair includes three verbs *look*, *see*, and *watch*. Based on the analysis of Wallwork (2018), *look* is mostly used with the preposition *at* and

generally means observe something with paying attention, while *see* is a verb of the senses and refers to one's physical sight. *Watch* is used for things that are moving, normally used with TV or movie. The differences of the words can be seen from (5a) to (5c).

(5a) Max turned to **look** at him. (preposition *at*, paying attention)

(5b) Nobody **saw** her when she fell down the stairs. (physical sight)

(5c) She loves to **watch** movies. (moving things)

Drawing upon these examples and analyses, it is anticipated that the participants will independently discern the distinctions between synonymous verb pairs through the raw corpus data.

3.4. Instruments

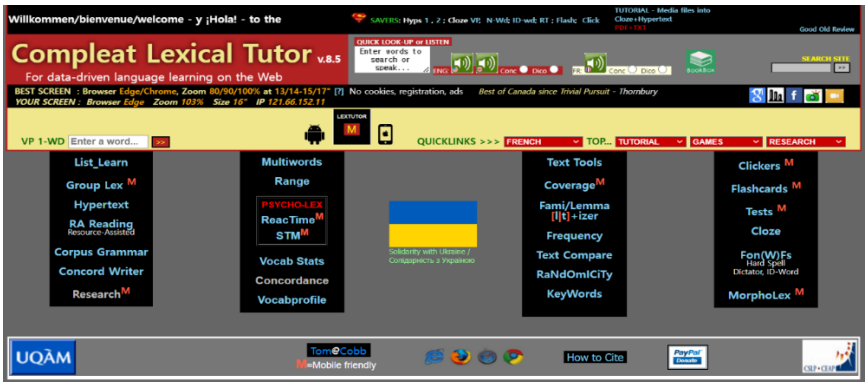
The present study employs relevant instruments to investigate the effects of computer-assisted DDL on distinguishing synonyms as stated in Chapter 1. In Section 3.2.1, the concordancing tool used in the current study is explained. Section 3.2.2 conveys the pretest and posttest materials utilized in this study. Section 3.2.3 describes the learning materials of computer-assisted DDL applied in the current study. Lastly, Section 3.2.4 presents questionnaire after the whole instruction.

3.4.1. Concordancing Tool

A concordancing tool used in the present study was "The Compleat Lexical Tutor" (*Lextutor*) version 8.5. *Lextutor* is a free website for data-driven learning developed by Tom Cobb based on his doctoral dissertation in 1997. *Lextutor* offers various corpus-based tools including concordance and vocabulary profile as shown in Figure 3.1.

Figure 3.1.

The Compleat Lexical Tutor v.8.5 interface[4]



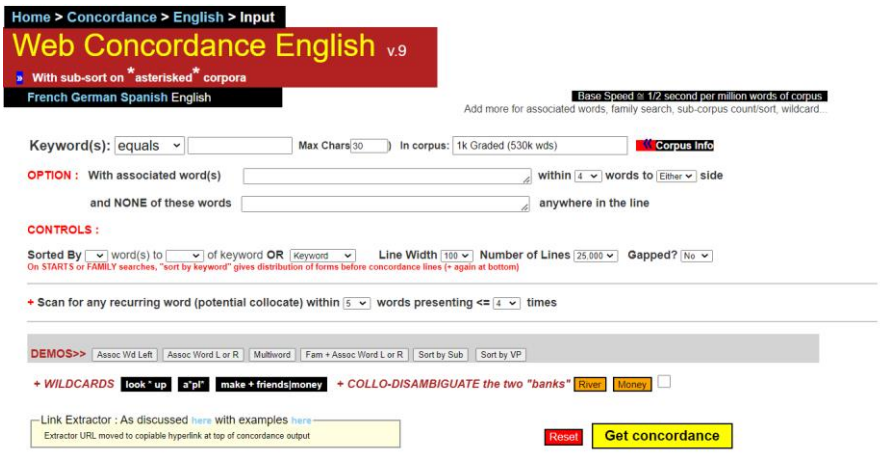
This study focused on using concordance, *Web Concordance English* version 9 (Figure 3.2). The concordance allows users to search keywords from the corpora provided by *Lextutor*. Moreover, users can select the options provided to

[4] Cobb, T. *Compleat Lexical Tutor* v.8.5 [computer program] Accessed 22 Oct 2022 at https://www.lextutor.ca

search certain keywords as follows: equals, starts, ends, contains, lemma and family. Within the scope of this study, the participants focused their attention on utilizing the ‘equals’ and ‘family’ options. The ‘equals’ option provides examples of target items that are an exact match, while the ‘family’ option encompasses a broad range of words associated with the target word, verb transformation form in particular. For example, when the participants enter the word ‘*look*’ with the ‘family’ option selected, they can get the concordance lines containing words such as *look, looks, looked,* and *looking*.

Figure 3.2.

Web Concordance English v.9 in Lextutor[5]



[5] Cobb, T. *Web Concordance English* v.9 [computer program] Accessed 22 Oct 2022 at https://www.lextutor.ca/con/eng

Among the diverse array of corpora available in this web-based tool, the 1K graded corpus (consisting of 530,000 words) was specifically selected for utilization. This particular corpus is tailored for beginners and represents a subset of the 2K graded corpus (comprising 1,200,000 words), which comprises numerous graded readers. These graded readers primarily consist of fictional narratives, alongside some non-fictional texts sourced from Oxford Bookworms, Penguin Readers, and Cambridge English Readers. Previous research by Geluso and Yamaguchi (2014) highlighted that learners often find interpreting concordance lines challenging during Data-Driven Learning (DDL) tasks. Consequently, Lee et al. (2018) suggested that an essential determinant of corpus effectiveness lies in the appropriateness of the corpus data provided to learners. Given that the 1K graded corpus is specifically designed for ESL beginners, it deliberately omits high-level vocabulary and adheres to the national curriculum of English for middle school 1st grade EFL learners in Korea. Thus, it is expected that the participants will be able to comprehend the contextual information within the concordance lines without requiring additional modifications from the instructor.

3.4.2. Pretest and Posttest Materials

The pretest material was made based on the selected five pairs of synonymous verbs. A total number of 20 questions included six questions requiring

the participants to choose an appropriate word, four for gap-fill questions with four options provided (Figure 3.3), six for correcting sentences, and four for gap-fill sentences (Figure 3.4).

(6a) ~ (6c) are some example questions of choosing an appropriate word. Students were requested to choose one word from the synonymous verb pair in order to make the given sentence complete and natural.

(6a) Perhaps they (heard / listened) me shout.

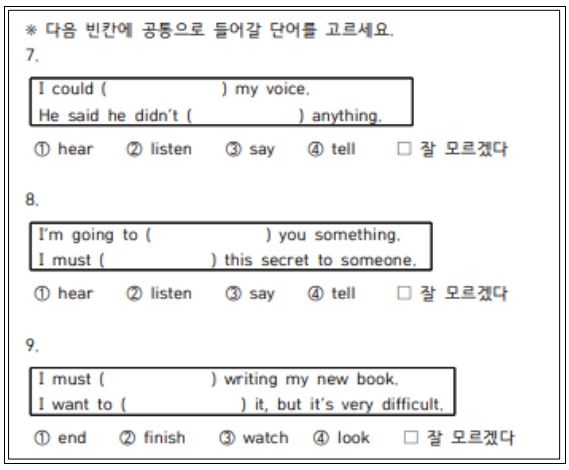
(6b) First, she had to (end / finish) her studies.

(6c) Nobody (looked / saw) her when she fell down the stairs.

Figure 3.3 presents multiple-choice gap-fill sentences. This type of question requires students to choose a common word which makes sense for the given two sentences.

Figure 3.3.

Test Example – Multiple-choice Gap-fill sentences



(7a) and (7b) below are examples for correcting sentences. Students were asked to check whether the words in the given sentences are appropriately used and make them correct and natural if not.

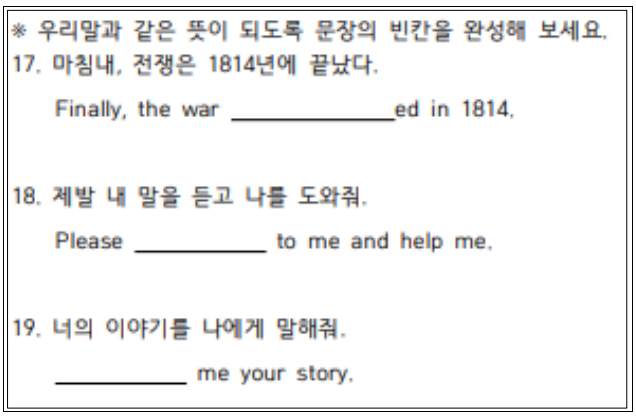
(7a) Anna listens carefully to her daughter.

(7b) They turned to watch at the house.

Lastly, Figure 3.4 shows some example sentences of gap-fill. In this type, students were expected to write down the correct word based on the given Korean translation of the sentence.

Figure 3.4.

Test Example – Gap-fill sentences



All sentences and contexts of test material were from *Lextutor*. Although the sentences and options were written in English, the instructions of the questions were provided in Korean, the participants' L1, to ensure the participants' understandings of instructions. In order to prevent participants' random guessing

when taking the test, all questions except fill-in-the-blank questions include the option of 'I'm not sure' (see Appendix 2 for details of pretest).

The posttest was conducted after the whole sessions of learning synonymous verbs. The types and number of questions in the posttest were identical with the pretest (see Appendix 3). However, the contexts and sentences were different from the pretest materials despite the use of the same concordancing tool, *Lextutor*. Among twenty questions, four questions with the highest error rate during the pretest were reselected from a set of twenty questions for the posttest, aiming to assess participants' ability to answer them correctly.

Table 3.2.

Reliability and Item Difficulty of the Pretest and Posttest

	Cronbach's alpha	Item Difficulty (*Mean*)	*N*
Pretest	.884	-.265	20
Posttest	.856	.946	19

Note. N=19 for the posttest as all participants got correct in one question.

The reliability and the item difficulty of the pretest and posttest were analyzed in advance. First, as the Table 3.2 shows, the Cronbach's alpha reliability of the pretest was .884 and the posttest was .856. As Gliem and Gliem (2003) stated, as an alpha of .8 considered as a reasonable goal of reliability, both pretest and

posttest can be regarded as reliable tests. However, one thing to notice is that all participants got the correct answer for one item in the posttest (Question number 2). Due to this 100% correct answer rate, this question was excluded when estimating Cronbach's alpha reliability of the posttest.

In addition to the reliability, the item difficulty was measured to statistically verify that posttest is not easier compared to pretest (see Table 3.2). The item difficulty of pretest with 20 items was calculated based on unidimensional one-parameter logistic IRT model. The average item difficulty of the pretest was -.265 which can be interpreted as having normal difficulty (between -.5 and .5). With fixing four common items and excluding one item showing perfect answer rate, the average item difficulty of the posttest was .946 (19 items) where over .5 represents that the questions are difficult. This result indicates that posttest was even more difficult than pretest. In conclusion, the prior analysis of pretest and posttest proved that both tests are reliable and the posttest was composed of relatively difficult items compared to that of pretest.

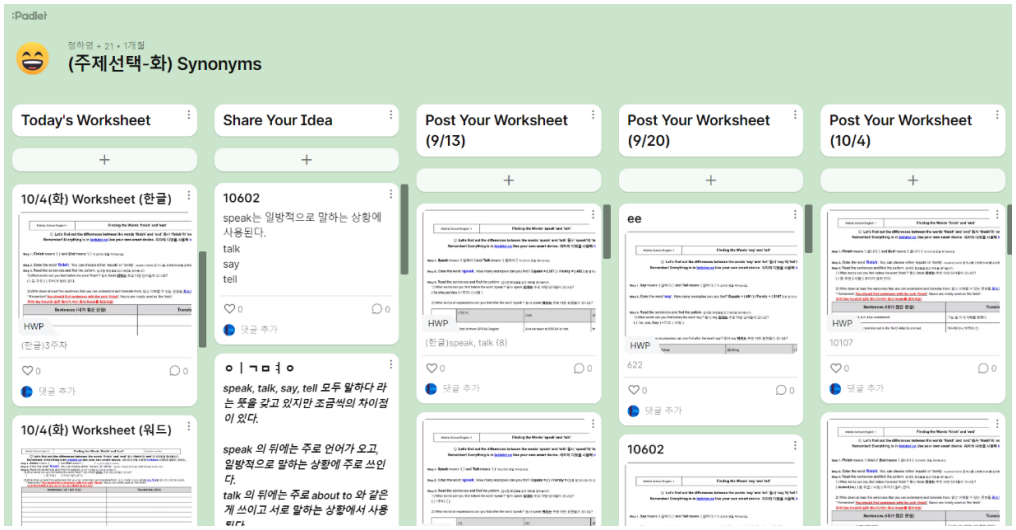
3.4.3. Learning Materials

The main learning material of this study is the digital device, Microsoft Surface Go 3. The participants used their own tablet PC during the whole session without any paper-based materials. To be specific, three kinds of learning materials

were used, which are *Lextutor*, *Padlet*, and worksheets. First, *Lextutor* is an online concordancing tool as previously explained in Section 3.2.2. The raw corpus data used in this current study were all from *Lextutor* so that each participant individual access to the website and search the target words every session. Second, *Padlet* (https://padlet.com), a collaborative web platform that participants can share their contents or ideas, is widely applied throughout the sessions (Figure 3.5). By using the dashboard offered by *Padlet*, all participants downloaded the worksheet, shared what kinds of differences they discovered regarding the target synonym pairs of each session, and posted their worksheet that they completed during each session.

Figure 3.5.

Padlet Dashboard Used Throughout the Sessions



The last learning material is the worksheet provided by the researcher every session using Microsoft Word. As the participants learned five pairs of synonymous verbs throughout three sessions, three different worksheets were provided to the students throughout the whole sessions (see Appendix 4 for a sample). The first worksheet for the first session dealt with the verb pair *speak* and *talk*, the second worksheet dealt with the verb pairs *say* and *tell*; *hea*r and *listen*, lastly the third worksheet dealt with the verb pairs *end* and *finish*; *look*, *see*, and *watch*. The form of the worksheet was basically the same for all synonymous verb pairs, which is based on the three-step procedure for concordance-based learning search suggested by Johns (1991) and the framework of corpus-based studies suggested by Flowerdew (2009). According to Johns (1991), the students first identify the target words and expressions then classify the characteristics and patterns of each target word based on the sentences they found followed by generalizing the usage of target words. In addition, Flowerdew's framework of corpus-based studies (2009) states 4 *I*'s; Illustration, Interaction, Intervention, and Induction. As some students might have difficulties discovering and inducing rules and patterns based on raw corpus data, intervention, the process of providing hints to students for an induction is included in the worksheet. In specific, there are some questions pertaining to the appropriate lexical and syntactic choices associated with a given verb in every worksheet.

Specifically, the worksheet basically includes six steps with the instruction written both in English and Korean (L1). First, the participants identify the target synonymous verbs by writing down the L1 meanings which are expected to be the same. Then, the participants search each target verb on web-based corpora and check how many examples exist. The third step is the main activity where the participants classify what expressions or words are found before and after each target verb. By reading the sentences they searched for, the participants classify the collocates or other contextual features of each verb and write down the discrete features and the following sentences on the worksheet. As this third step is significant for differentiating synonymous verbs, the researcher's intervention is incorporated as a guiding prompt for certain word pairs, particularly in cases where participants encountered challenges in identifying distinctions in contextual meaning. Furthermore, certain synonymous pairs necessitate participants' accurate comprehension of the connotations conveyed in the concordance lines, as presented in Table 3.1. Consequently, in the third step of such pairs, a translation activity is incorporated. Participants are tasked with translating the identified concordance lines into Korean, thereby endeavoring to grasp the contextual usage of the words.

The next step is generalizing what participants have discovered. Based on the concordance lines of each verb pair, the participants are expected to find out the similarities and differences between the verb pair. This step is followed by fill-in-

the-blank activity where participants check their knowledge of discriminating synonymous verbs. The last step is making new sentences with the exact use of synonymous verbs they have learned. By completing the worksheet, the participants are encouraged to discover the differences between the synonym pairs and use them in the correct context.

3.4.4. Questionnaire

A questionnaire developed for this study was conducted after the whole session and the posttest. This questionnaire is aimed to examine the participants' perception of computer-assisted DDL (see Appendix 1) based on Likert-scale. To investigate learners' attitudes and perceptions toward DDL and CALL in detail, ten questions (Q1 ~ Q10) were asked about DDL using raw corpus data during class and other five questions (Q11 ~ Q15) dealt with CALL as can be seen from Table 3.3. Among ten questions dealing with DDL and corpora, seven questions (Q1, Q3 ~ 7, Q10) were taken and revised from Hong and Oh (2008), and one question (Q8) was written based on Park's study (2008). The remaining two questions (Q2, Q9) were additionally designed and surveyed for this study, focusing on advantage of using corpora and learning autonomy. Lastly, five questions about CALL were taken and revised from Talebinezhad and Abarghoui (2013) to examine learners' attitudes and any possible difficulties of using computer during class.

Table 3.3.

Likert-Scaled Survey Item List

	Survey Item
Q1	Using corpora helps me learn the differences between synonyms.
Q2	Corpora is helpful to learn the usage of words in the sentences.
Q3	Using corpora provides me opportunities to learn authentic English.
Q4	It took a lot of time to use and analyze corpus data.
Q5	I had trouble because I didn't know may words in corpora.
Q6	I had trouble because there were too many sentences in corpus data.
Q7	Overall, corpora are useful learning materials.
Q8	I'd like to keep using corpora to do other activities in English classes.
Q9	I can study English by myself using corpora.
Q10	I'd like to recommend using corpora in other English classes.
Q11	CALL is interesting and stress-free environment to learn English.
Q12	CALL helps me identify and discriminate synonymous verbs.
Q13	I do not know how to make use of computers so the lesson was difficult.
Q14	It took less time to finish the worksheet because of using computer.
Q15	I'd like to keep using computer in English classes.

The participants were prompted to answer using a 5-point Likert scale ranging from 1 ('strongly disagree') to 5 ('strongly agree'). Additionally, three open-ended questions are included to investigate learners' thought of DDL and CALL: 1) What was the biggest advantage of learning English using corpus data and computer? 2) What was the biggest difficulty of learning English using corpus

data and computer? 3) Please comment about the whole session if you have any other ideas or opinions. Cronbach's alpha reliability of the questionnaire was .896, which is proved to be reliable.

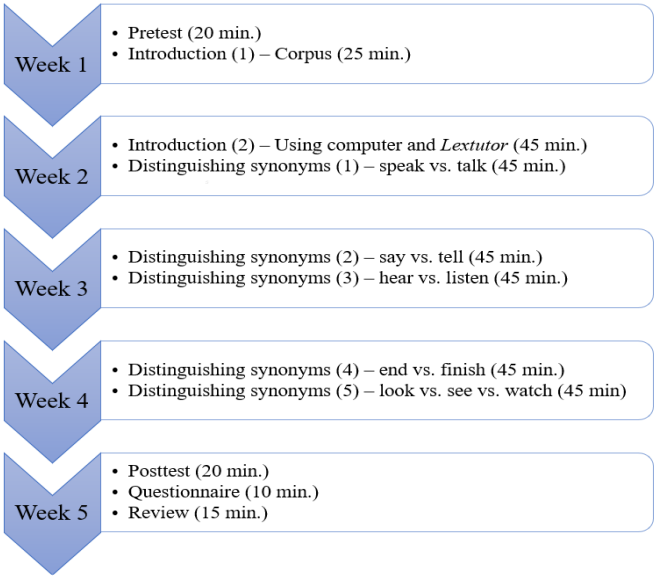
As some participants might have difficulties interpreting questions due to their English proficiency, all questions were translated into Korean and presented to the participants.

3.5. Procedures

The experiment of this study lasted for three weeks with ninety minutes class each week, and the additional pretest and posttest session for one week each. Although a class duration of middle school in Korea is forty-five minutes, free year program classes are originally designed as ninety minutes long, one class per week. The participants got ten minutes break time after the first session of forty-five minutes. The whole procedures were conducted in English classroom and the participants brought their own digital device every week except the first week of the experiment. The researcher took on the role of an instructor and the whole session was based on individual learning activities. The summary of the whole experiment procedure is illustrated in Figure 3.6.

Figure 3.6.

Summary of Experiment procedure



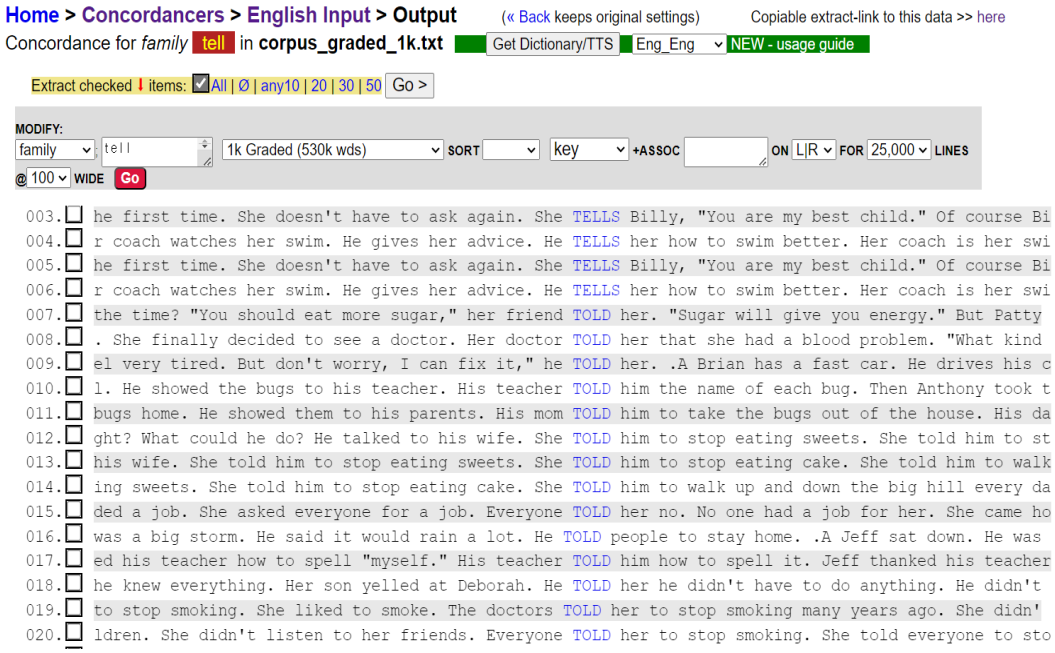
On the first week, the experiment started from the latter half of the session (forty-five minutes). All participants took the pretest for twenty minutes. Then, the instructor introduced the concept of corpus to the participants in Korean for twenty-five minutes and helped them understand the purpose of using corpora in our experiment.

From the second week, the experiment took the whole ninety minutes of class time. The training session was conducted on the first session of the second week class as it is important to provide appropriate DDL training to pre-tertiary learners (Papaioannou et al., 2020; Liontou, 2020). The training session lasted for forty-five minutes. As the current study focuses both on DDL and CALL, the

participants practiced how to find words using *Lextutor* along with how to organize and enter data using computer (see Figure 3.7 for a result in *Lextutor*).

Figure 3.7.

Sample Result of Verb 'tell' in Lextutor



Then, the main activity was conducted from the latter half of the session. All participants first downloaded the worksheet from *Padlet*, accessed to *Lextutor* and entered the target word. Following the steps written in the worksheet which are described in Section 3.2.4, the participants individually worked to distinguish synonymous verb pair. However, as the participants all have different English

proficiency level, the instructor's intervention was included based on Flowerdew (2009) especially for low and intermediate students. After the participants discovered the patterns and usage of synonymous verb pair for about thirty minutes, they were encouraged to share their ideas on *Padlet*. Based on those ideas, the instructor consolidated what participants discovered and asked them to submit the worksheet on *Padlet* (see Table 3.4 for the detailed lesson procedure).

Table 3.4.

Lesson Procedure

Step	Learning Activities	Time Allotment
Introduction	- Instructor introduces the target verbs - Learners download the worksheet from *Padlet*	5 min
Development	[Identify] - Learners enter the target verbs on *Lextutor* - Learners read the concordance lines	6 min
	[Classify] - Learners classify the sentences by structure or meaning	12 min
	[Generalize] - Learners generalize the usage of synonymous verbs (Instructors' intervention provided if necessary) - Learners explain the differences between the verbs - Learners share their ideas on *Padlet*	12 min
Consolidation	- Instructor reviews the learners' findings - Learners complete the worksheet and submit on *Padlet*	10 min

The sessions of week three and four was identical with week two. However, as some verb pairs (hear, listen; end, finish; look, see, watch) required learners to infer differences in contextual meaning, the instructor's intervention became

significant. As a kind of intervention, the instructor led the participants' discovery by providing specific sentences from *Lextutor*. On the last week (Week 5), all participants took the posttest for twenty minutes followed by the questionnaire for ten minutes. After all the experiment ended, the instructor additionally conducted a review session about synonymous verbs for fifteen minutes

The interview with three participants were conducted from the second week to the last week as described in Section 3.2.6. They individually had an interview with the instructor after each week's class ends. The interview was also held in English classroom with the participants' own digital device, watching the screen recording together.

3.6. Data Analysis

The present study implements mixed-methods research to examine three research questions described in Chapter 1. This section introduces statistical data analysis along with qualitative data analysis according to the research questions. As Table 3.5 shows, quantitative approach was applied to examine the first and second research question, while qualitative approach was applied to examine the last research question.

Table 3.5.

Research Questions and Data Analysis Methods

Research Questions	Data Analysis Methods
1. Do Korean EFL middle school students improve their comprehension ability to distinguish synonymous verbs through computer-assisted DDL?	Quantitative Method (Pretest, Posttest)
2. How do learners perceive computer-assisted DDL?	Quantitative Method (Questionnaire)
3. What changes do learners experience their cognitive and affective domains through computer-assisted DDL?	Qualitative Method (Interview, Screen Recording)

3.6.1. Quantitative Approach

The first research question of the present study was 'Do Korean EFL middle school students improve their ability to distinguish synonymous verbs through computer-assisted DDL?' In order to answer this question, the participants' pretest and posttest scores were compared and analyzed. The main statistical analyses were carried out using the statistic software *SPSS ver. 25* and *flexMIRT ver. 3.6.5*. First, the reliability of two tests were examined based on Cronbach's alpha coefficient. Then, by using *flexMIRT ver. 3.6.5*, the item difficulty of the pretest and posttest were analyzed to verify that overall difficulty of the two different tests based on IRT (Item Response Theory) are not significantly different. According to Seong (2016), item difficulty less than -.5 represents easy question and coefficient over .5

represents difficult question. Item difficulty between -.5 and .5 is considered as a normal question. After verifying the reliability and item difficulty of these two tests, each participant's pretest and posttest scores were compared to see the effect of computer-assisted DDL on distinguishing synonymous verbs. Due to relatively small number of participants (n=29), a non-parametric statistics for paired sample t-test was employed with the level of significance set at .05.

In addition, the effect size was calculated in the results of the t-test. Although Cohen's (1988) labels for the effect size as small (d = 0.2), medium (d = 0.5) and large (d = 0.8) have been widely used in research, Plonsky and Oswald (2014) argued that L2 researchers comparing pre and post groups should consider the scale as small (d = 0.6), medium (d = 1.0) and large (d = 1.4) to interpret the importance of L2 research effects more accurately. Therefore, this study will also follow the suggested scale when interpreting the effect size.

The second research question was 'How do learners perceive computer-assisted DDL?' As previously explained in Section 3.2.5, the fifteen Likert scale-based questions and three open-ended questions in the questionnaire were asked to the participants to analyze the participants' perception and attitude toward computer-assisted DDL. Cronbach's alpha coefficient was checked in advance to guarantee the reliability of this questionnaire. Among fifteen Likert scale-based questions, there were eleven positively-worded questions (Q1 ~ Q3, Q7 ~ Q12, Q14

~ Q15) and four negatively-worded questions (Q4 ~ Q6, Q13). For all questions, ‘strongly agree’ was coded to 5 while ‘strongly disagree’ was coded to 1. Three open-ended questions were examined and organized based on the emerging themes. This analysis is expected to expand the attitudes and perceptions of EFL Korean middle school students on computer-assisted DDL by supporting the results of the quantitative analysis of questionnaire.

3.6.2. Qualitative Approach

To examine the third research question, ‘What changes do learners experience through computer-assisted DDL?’ the interview and screen recording of three participants were conducted throughout the experiment.

First, the interview was conducted to scrutinize participants’ thoughts and feelings about computer-assisted DDL along with their process of discovering similarities and differences between synonymous verb pairs. Among some participants who volunteered to have an interview, the researcher selected three of them showing obvious differences in their English learning background and proficiency.

The researcher interviewed three participants after each session (four times in total) for about ten to fifteen minutes. As the interview was a semi-structured interview, the students were asked pre-selected questions (see Appendix 5) and

based on their response and reaction the researcher added questions for a deep understanding of students' attitude. All interviews were conducted in Korean as every participant's L1 is Korean. In addition, for the qualitative analysis of interview, all interviews were recorded.

In the process of analyzing interview data, the voice recordings of the participants' responses were transcribed by the researcher. As the main goal of transcribing the interview was to organize the emerging themes related to each session and observe each participant's progress or difficulties, their intonation and nonverbal utterances were not included in the transcriptions.

Second, while the participants completed their worksheet during the session, the entire computer display was recorded using a software called *oCam* (https://ohsoft.net). Among various screen recording programs, this software was the most appropriate in Microsoft Surface Go 3 and easy to use for pre-tertiary participants. The participants clicked the recording button before starting every session and clicked the stop button after they finished their worksheet to save the recording file. In fact, screen recording was applied in the present study to analyze qualitative data deeply. The researcher and the participant watched the recorded video together during the individual interview. This instrument makes it possible for the researcher to ask some specific questions about the process of discriminating synonymous verbs and investigate learners' learning behavior.

For the analysis of screen recording data, the researcher observed the recording and described the screen with the participant's action. For example, if the participant scrolled down the screen rapidly, the researcher wrote the possibility of paying less attention. If the participants' mouse cursors moved along the sentences, the researcher wrote the possibility of reading the sentences. In addition, a time stamp is added as it played a significant role in the qualitative analysis. This is because time spent in one window (i.e., concordance lines in *Lextutor* or worksheet) for a long time could be interpreted as difficulties with the task completion. Based on the observation, the researcher organized the action timeline each session. A short excerpt of an action timeline is illustrated in Figure 3.8.

Figure 3.8.

Sample Action Timeline

Time	Screen	Actions
00:08-01:00	Padlet	Accessed to the Padlet and downloaded the worksheet
01:00-01:37	Word document (Worksheet)	Scrolled down the screen to read the worksheet
01:37-02:25	Lextutor	Interface - Concordance - English - set 'equals' and corpus as '1k Graded'
02:25-03:10	Worksheet	Entered the definition of target word in L1 (Step 1)
03:11-03:54	Lextutor	Entered the word *speak* and scrolled the screen slowly (reading concordancing lines)

Based on the transcriptions and the action timeline of the three participants, their specific behavior and attitudes toward the computer-assisted DDL are described in the following chapter.

CHAPTER 4.
RESULTS

This chapter analyzes the quantitative and qualitative research data collected during the experiment. To present the result, the research questions stated in Chapter 1 are answered consequently. Section 4.1 addresses the first research question on whether computer-assisted DDL enhances EFL middle school students' ability to distinguish synonymous verbs. The quantitative data of pretest and posttest are analyzed to clarify the effectiveness of computer-assisted DDL. Next, Section 4.2 answers the second research question which aims to investigate learners' perception toward computer-assisted DDL based on the questionnaire. Lastly, Section 4.3 deals with the analysis of learners' detailed behavior during the experiment to examine the changes they show throughout the session, which is supported by qualitative data including individual interview and the screen recording.

4.1. Effects of Computer-assisted DDL on Distinguishing Synonymous Verbs

In order to examine the effects of computer-assisted DDL on distinguishing synonymous verbs, the first research question of this study, the participants' test

scores were compared. Table 4.1 summarizes the descriptive statistics retrieved from the results of pretest and posttest with mean, standard deviation, minimum and maximum scores of the tests. As each test had twenty questions, both tests had a total point of 20.

Table 4.1.

Descriptive Statistics of the Pretest and the Posttest (N=29)

Test	*Mean*	*SD*	*Min*	*Max*
Pretest	10.55	5.14	0	19
Posttest	14.00	4.40	4	20

Based on the descriptive statistics, mean score of the posttest is higher than that of the pretest as it increases from 10.55 (*SD* = 5.14) to 14.00 (*SD* = 4.40). For the detailed analysis, 12 out of 29 students got wrong for more than half of the whole items in the pretest. However, the result of the posttest shows a sharp decrease of this range to 4 students in total (see Appendix 6 for the whole pretest and posttest scores of the participants).

In addition to the descriptive analysis, a paired sample t-test was conducted to examine whether the means of the pretest and posttest of synonymous verbs using computer-assisted DDL showed a statistically significant

difference. Table 4.2 illustrates the overall results of the paired sample t-test and the effect size of it.

Table 4.2.

Results of a Paired Sample t-test of the Pretest and the Posttest

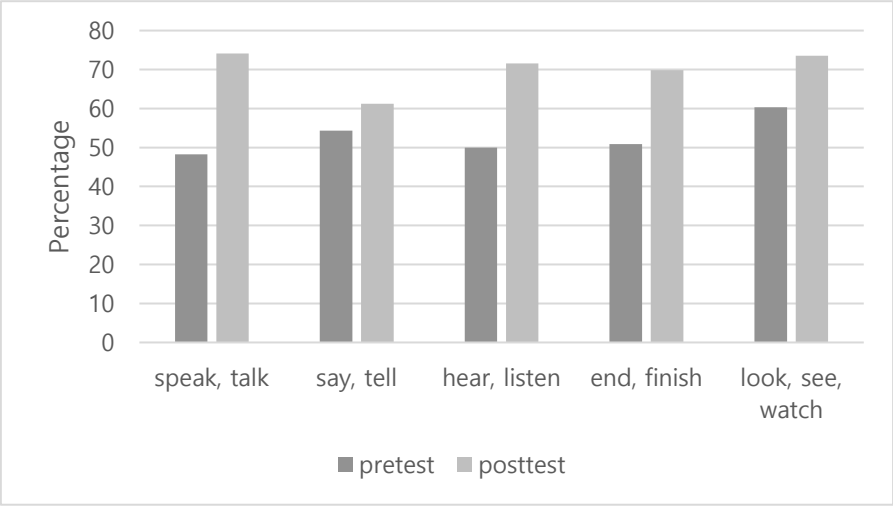
	Paired Differences					*t*	*df*	*Sig.*	*Effect size*
	Mean	*SD*	*SE*	95% CI					
				Lower	Upper				
Post-Pre test	3.45	3.28	.61	2.20	4.70	5.66	28	.000	1.05

The analysis of the paired sample t-test as in Table 4.2 demonstrated that there is a significant difference between the pretest and the posttest scores regarding synonymous verbs (t = 5.66, p < .001). Moreover, the effect size is considered as medium (d = 1.05) based on the criteria suggested by Plonsky and Oswald (2014)[6]. This analysis suggests that computer-assisted DDL lessons on distinguishing synonymous verbs have medium positive impacts on increasing learners' ability to distinguish the words. In addition, unidimensional one-parameter logistic IRT model demonstrated a mean growth pattern of the participants as 1.98, which can be interpreted as high growth.

[6] As previously mentioned in Section 3.4.1, L2 researchers comparing pre and post groups should consider the effect size as small (d = .6), medium (d = 1.0) and large (d = 1.4) to interpret the importance of L2 research effects more accurately.

Figure 4.1.

Percentage of the Correct Answer of Each Synonymous Verb Pair



To be specific, the percentage of correct answer of each synonymous verb pair all increased as Figure 4.1 shows. In the pretest, the participants were most likely to have incorrect answers of the pair [*speak, talk*] (48.28%), followed by [*hear, listen*] (50%) and [*end, finish*] (50.86%). The participants relatively got more correct answers on [*say, tell*] (54.31%) and [*look, see, watch*] (60.35%) pairs. The result changed in the posttest as [*speak, talk*] showed the highest percentage of correct answer (74.14%) followed by [*look, see, watch*] (73.56%), [*hear, listen*] (71.55%) and [*end, finish*] (69.83%). The pair [*say, tell*] relatively increased less than other pairs (61.21%). One possible explanation might be the difficulty of distinguishing the concept of message-focused and addressee-focused compared to the other concepts.

In conclusion, despite some small differences between each synonym pair, the participants showed growth of their ability in average even though the item difficulty of the posttest becomes higher. The statistical result of the pretest and the posttest proves the positive effects of computer-assisted DDL on distinguishing synonymous verbs in case of middle school EFL learners.

4.2. Learners' Perception toward Computer-assisted DDL

The fifteen 5-point Likert scale questions and three open-ended questions in the questionnaire were analyzed to examine learners' perception toward computer-assisted DDL, the second research question. The result of the fifteen 5-point Likert scale questions is summarized in Table 4.3 with mean score and standard deviation of 29 participants.

As the result shows, the participants generally perceive computer-assisted DDL positively, with mean scores of 5-point Likert scale questions distributed from 3.52 to 4.24, except four negative questions (Q4, Q5, Q6, Q13). In case of standard deviation, the result distributed from .69 to 1.64. Question 15 showed the highest deviation which means that the participants' response toward keep using computer in English classes varied a lot. Some students preferred using computer every class while others did not. For the detailed analysis, the questionnaire can be divided into DDL and CALL field respectively.

Table 4.3.

Results of the Questionnaire

Questions	*Mean*	*SD*
Q1. Using corpora helps me learn the differences between synonyms.	4.00	1.00
Q2. Corpora is helpful to learn the usage of words in the sentences.	4.03	.96
Q3. Using corpora provides me opportunities to learn authentic English.	4.21	.88
Q4. It took a lot of time to use and analyze corpus data.	2.31	1.01
Q5. I had trouble because I didn't know may words in corpora.	2.03	.89
Q6. I had trouble because there were too many sentences in corpus data.	2.31	1.44
Q7. Overall, corpora are useful learning materials.	4.00	.79
Q8. I'd like to keep using corpora to do other activities in English classes.	3.52	1.12
Q9. I can study English by myself using corpora.	3.69	1.08
Q10. I'd like to recommend using corpora in other English classes.	3.86	.69
Q11. CALL is interesting and stress-free environment to learn English.	4.24	.90
Q12. CALL helps me identify and discriminate synonymous verbs.	4.03	1.03
Q13. I do not know how to make use of computers so the lesson was difficult.	2.03	1.61
Q14. It took less time to finish the worksheet because of using computer.	3.72	1.14
Q15. I'd like to keep using computer in English classes.	3.72	1.64

Table 4.4 indicates that the participants answered CALL related questions slightly more positively compared to DDL related questions. Among all questions, Question 11 asking whether CALL is interesting way of learning showed the highest mean score (*Mean* = 4.24). In addition, the participants responded that CALL is useful (Q12, Q13, Q14) so that they want to keep using computer in English classes (Q15). Similarly, the participants generally showed positive perception and attitude toward DDL. Most prominently, the participants agreed that corpora provide "authentic" examples of English, showing the second highest mean score (*Mean* = 4.21) in Question 3. Moreover, many participants expressed corpora as a useful tool of learning English (Q2, Q7) including synonyms (Q1). However, some participants revealed the difficulty of using corpora during class because corpora contain considerable number of examples (Q6) and it took a lot of time to analyze corpus data (Q4).

Table 4.4.

Results of the Questionnaire (DDL / CALL with 5-point Likert scale)

	N	*Mean*	*SD*
DDL questions	10	3.87	.21
CALL questions	5	3.94	.22

The results of three open-ended questions were in parallel with 5-point Likert scale questions. The participants were recommended to write some positive and negative aspects of the whole lesson including free last comment. All participants responded to the question asking positive aspects but 18 out of 29 participants wrote negative aspects.

Table 4.5.

Learners' Overall Reactions Toward Computer-assisted DDL

Reaction Type	Main Responses (Number of participants)
Positive	1. Corpus was a useful tool to learn synonyms and words. (*N*=11) 2. Accessing authentic example sentences was interesting and helpful. (*N*=6) 3. Computer-assisted DDL was fun and interesting. (*N*=4) 4. Computer-assisted DDL helped me to participate actively on class. (*N*=3) 5. It was easy to finish the work by using computer. (*N*=5)
Negative	1. Using computer was complex. (*N*=8) 2. It was difficult for me to understand the sentences in the corpus data. (*N*=5) 3. It was difficult for me to discover the differences between synonyms. (*N*=5)

Table 4.5 summarizes the learners' overall reactions toward computer-assisted DDL lesson of distinguishing synonymous verbs. First, 35% (*N*=11) of the participants wrote about the usefulness of corpus as a vocabulary learning tool. In specific, the answers varied including "Corpus helped me to understand the clear differences between the synonyms" and "I was able to get the sense of what

expressions go with certain words". Secondly, the participants mentioned reading many authentic examples in corpora. For example, about 22% (*N*=6) of the participants responded that "It was helpful because I was able to read many real usages of certain words", "Reading authentic examples was interesting and I felt like I was learning real English". Some participants also wrote about their learning attitude that they become an active learner throughout the lessons. Moreover, they answered that by searching the words themselves, they could understand and remember what they learned better compared to their normal English lessons.

In contrast, negative reactions were concentrated in using computer. 8 participants mentioned that using computer was complex and difficult for them. Other negative reactions were about the difficulty of understanding sentences and discovering the differences between synonyms themselves. To conclude, the questionnaire indicated that the participants perceived computer-assisted DDL positively as the corpus provided lots of authentic examples, helping learners to successfully understand and distinguish synonyms. However, it is obvious that some participants showed negative perception toward computer-assisted DDL because first, they are not used to operating computer for learning and second, inferring rules based on corpus data is difficult to some of them.

4.3. Descriptive Accounts of Learners' Behavior and Responses

To examine participants' changes in behavior and attitude throughout the session, the third research question of this study, the individual interviews that were conducted with three participants are analyzed in this section. As previously mentioned in Section 3.2.6, the researcher had a semi-structured interview with three participants after each session, four times in total.

Table 4.6.

The Pretest and the Posttest Score of Three Participants

	Pretest Score	Posttest Score
Hanna	17	19
Irene	11	16
Leo	4	14

As Table 4.6 shows, Hanna is considered as a high proficiency learner in this study as she got 17 out of 20 in her pretest. After the whole session, her score become 19 out of 20. Irene's pretest score was 11, which is considered as an intermediate proficiency learner and her posttest score became 16. The last participant, Leo is a low proficiency learner of English as he got only 4 out of 20 in

his pretest. However, after the whole session, his score drastically increased into 14.

The descriptive accounts of each participant's behavior and responses are presented in the following subsections. First, the participant's performance on tasks is summarized. As there were three sessions, each participant's change and development throughout the sessions are demonstrated. Then, the participant's performance on using computer is described followed by each participant's overall impression on computer-assisted DDL. The descriptive accounts are provided based on both transcription of interview and screen recording.

4.3.1. Hanna: An Advanced Learner with Active Participation

Hanna showed a great performance throughout the whole session. She reported as in Excerpt 4.1 that she likes English and she thinks that she is good at English. She has heard about corpus when she was an elementary school student. Specifically, it was an afterschool computer class and she remembered that her teacher explained briefly about the concept of corpus. She recalled:

> Excerpt 4.1. I remembered the word 'corpus' but I think the teacher did not explain it in detail like today's class. The teacher just described corpus as a data collection with people's spoken and written words. I have never actually used corpus. It is my first time using it.[7] (Hanna's Interview Quote,

[7] All of the verbal protocol excerpts presented in this chapter were translated from

Session 1)

As Hanna had background information about the corpus, she adapted to the DDL lesson in a short time and actively participated in the whole session. In the following subsections, her performance on tasks, using computer, and her impression on DDL are described in detail.

4.3.1.1. Performance on Tasks

Hanna finished her task on distinguishing synonymous verbs on time every session. On her first session, searching each verb *speak* and *talk*, and discovering the similarities or difficulties between those words took about 38 minutes in total without any error. Hanna reported in Excerpt 4.2 that it was not difficult for her to use corpus because of the practice session, a detailed explanation about corpus and the way of using it. Based on what she practiced, she followed the direction on the worksheet to distinguish synonym pairs. However, when she was finding the first word *speak*, there was some interval between each direction.

> Excerpt 4.2. At first, I tried to understand how this lesson is going on. So I waited until the teacher explained about it. After finding the first word, *speak*, I could fully understand the whole process. From the next word,

Korean to English by the researcher.

talk, I finished the worksheet on my own and it was not difficult. (Hanna's Interview Quote, Session 1)

From then, Hanna quickly adjusted to the lesson procedure. She responded that she did not exactly know the actual usage of certain words and the differences between synonymous verbs before, but as she used corpus and found the words herself, she could understand the words better than before. Moreover, she replied that DDL tasks helped her to fully internalize the knowledge she learned during the session.

On the second week, Hanna completed the task by herself without any support from the teacher. To differentiate the words *say* and *tell*, she accessed the corpus, entered each word and read the expressions following and followed by the target word as the target word is marked in blue. Then, she scanned through the sentences again so that she could find the common usage of the word. She reported as in Excerpt 4.3 that the whole process was not difficult so she could finish it early.

Excerpt 4.3. I felt like I got used to doing this task. As I perfectly know what and how to do it, I could find everything by myself. For the words *say* and *tell*, the characteristics were obvious for me so I could finish it little bit earlier this time. I first entered what I found in my worksheet and compared my answer with the teacher's hint later. Then, I finally

organized the similarities or differences on my worksheet. (Hanna's Interview Quote, Session 2)

Interestingly, Hanna also reviewed what she found last time to summarize all four 'speak' category words – *speak, talk, say, tell*. She opened her last week's worksheet and reviewed through it to retrieve what she found and wrote. After the self-directed review, she summarized and synthesized all the information she found based on the corpus data. She later reported through the interview session that she could differentiate all four 'speak' category words and apply those words into new sentences as she read many authentic examples in corpus data.

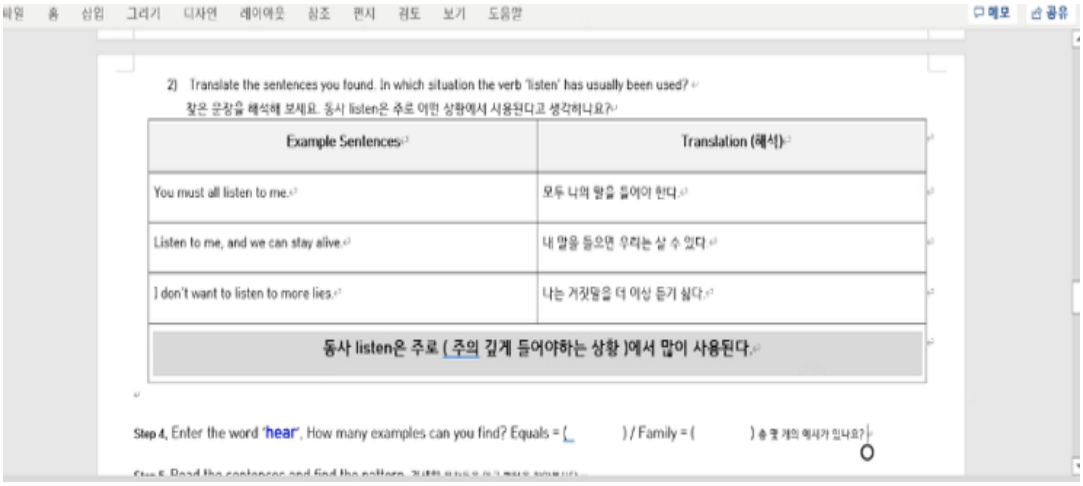
The following pairs of synonymous verbs required participants to translate the sentences in concordances to identify the context that certain words are used. Hanna emphasized that she likes translating the concordance lines as it helped her a lot (see Figure 4.2). In detail, Hanna recalled:

> Excerpt 4.4. I really liked the part that includes translating. That was very helpful. In fact, I didn't read the sentences thoroughly while searching the words in corpus before. However, today, I tried to read the sentences carefully and figure out the context that word is usually used. I think because of the translation, I could discover the differences between the synonyms which are not obvious. Also, I think I could improve my reading and translation ability while doing this activity so I prefer doing

translation. (Hanna's Interview Quote, Session 2)

Figure 4.2.

Screenshot from Hanna's Worksheet



Hanna, a high proficiency learner of English, repeatedly stated that she preferred translating as it is more challenging and interesting to her. In addition, she reported the last task, comparing three words - *look, see, watch* - at once, was the most impressive because it required her to think a lot. She answered that she was able to concentrate throughout the whole class time because she aimed to discover the similarities and differences among three words herself. In conclusion, Hanna's performance and responses indicate that she was able to control the learning process herself, becoming an autonomous learner.

4.3.1.2. Performance on Using Computer

Hanna did not show any difficulties of using computer throughout the whole session. She replied that she usually uses tablet PC when she studies at home so she prefers writing on computer rather than writing on paper. She added another comment that as each lesson required searching words online and writing down the sentences, completing the worksheet with computer seems to be much more reasonable.

The only problem of using computer that Hanna mentioned is the WiFi access. In the first interview, Hanna recalled:

> Excerpt 4.5. Completing the worksheet using computer was not difficult for me. However, my computer did not work well today. I think it is because of the WiFi access in school. I had to wait for a minute to get the result when I entered a word on the corpus website. It would be much better if my computer and WiFi works fast with a short runtime. (Hanna's Interview Quote, Session 1)

As Hanna responded, some intervals were recorded during the lesson. She waited until she could fully access to the website and then continued her task. However, the problem did not last for a long time. On the second and third week,

Hanna mentioned that her computer was not a problem at all. Compared to the first week, she agreed that WiFi worked well so that she could concentrated on DDL task itself.

4.3.1.3. Perception on Computer-assisted DDL

Hanna was requested to describe her perception on computer-assisted DDL at the last interview. She provided some advantages and disadvantages she has thought about it. Hanna reported that computer-assisted DDL has only advantages in her aspect. At first, she mentioned that computer-assisted DDL was an effective way to learn how to distinguish synonymous verbs. Although she did not have any chances to consider the different usages of synonyms before, this whole session encouraged her to differentiate each pair of synonyms. Moreover, Hanna emphasized her preference of studying and learning herself as an autonomous learner. She recalled:

> Excerpt 4.6. Compared to normal English classes in school, this lesson was much more interesting to me. We just memorized the definition of English words in Korean, which was usually boring. However, as I found example sentences myself by using computer and corpus data, I could concentrate on the class every time and could remember the information I discovered very naturally for a long time. I felt like I am managing and

> controlling my learning process very well. I want to learn English with this kind of methodology every time. (Hanna's Interview Quote, Session 4)

In addition, Hanna volunteered to search other words on her own. As she got used to *Lextutor* website, she entered some difficult vocabularies she learned in the private institute and read the authentic examples of the word. Also, she asked the researcher whether there are other synonymous verb pairs that she could study herself.

Regarding the aspect of CALL, Hanna responded using computer during class itself is interesting so that students including herself tended to participate actively on class. Also, as their educational device allows them to search through Internet, they can find out what they do not know on their own and complete the worksheet despite the level differences. When some minor logistic problems such as WiFi connection are solved, Hanna totally agreed on using computer in English classes.

4.3.2. Irene: A Hardworking Intermediate Learner

The next participant, Irene, is a female participant who has an intermediate English proficiency level. Although she is not a fluent English speaker and she has never heard about corpus before, she worked hard throughout

the session to complete her tasks. A detailed behavior of Irene based on the interview and screen recording is described in the following sections.

4.3.2.1. Performance on Tasks

In the first interview, Irene showed curiosity toward the whole lesson procedure. She replied that searching the words using *Lextutor* website and completing the worksheet using computer was totally new to her. However, her performance on tasks proved that she had some difficulties on distinguishing synonyms based on corpus data. For the first word pair, *speak* and *talk*, it took the whole 45 minutes to complete the worksheet. In addition, she recalled in Excerpt 4.7 that she could not easily figure out the common usage of certain words. The screen recording proved this as Irene's screen frequently stopped for a minute without showing any movement. Irene said:

> Excerpt 4.7. At first, it took some time for me to understand and get used to the whole procedure. For me, to find a common usage of the word *speak* was difficult so I waited until the teacher explains it. I first entered the word *speak* and scanned through the example sentences but I could not categorize the common usage of the word. After the teacher gave us a hint that words such as Chinese, English, Spanish come with *speak*, I was able to write the example sentences in my worksheet. (Irene's Interview Quote,

Session 1)

As Irene reported, it took some time for her to understand how to find similarities and differences between the synonyms. She could easily figure out the characteristics without hesitation which dominate the concordances such as ‘talk to’, but she expressed difficulties when the characteristics are not obvious.

In the second week, Irene was requested to describe what she felt about the lesson compared to the first week. She responded that she becomes proficient as she completely understands what the teacher intends and what she should do. In fact, as the class started, she immediately downloaded the worksheet, accessed *Lextutor* website and entered the first word *say* by herself. Also, the interval or hesitation during the lesson decreased. Although she got used to the task on distinguishing synonymous verbs, she said she still needed some hints or supports to check that what she discovered was correct. After finishing the second task, distinguishing *say* and *tell*, the participants were asked to compare four ‘speak’ category words – *speak, talk, say, tell* – in total. Irene recalled this activity and reported as in Excerpt 4.8 that she could not remember what she figured out last week at first. She explained what she did in detail:

> Excerpt 4.8. Honestly, I could not remember what I found last week, about *speak* and *talk*. So I hesitated for a moment, but I suddenly realized that I

> have a computer to find the differences again. Then, I entered the word *speak* and *talk* again in *Lextutor* website. The amazing thing was that when I read only a few sentences, I could remember what I learned last week. This is how I could complete the last activity. (Irene's Interview Quote, Session 2)

Despite Irene's limited proficiency level, she managed to overcome the challenge of not recalling what she had learned the previous week. Through her own efforts, she found a solution and independently summarized the distinctions among the four words belonging to the 'speak' category. She additionally commented that it was a fresh experience because she normally waits until the teacher gives the answer.

Interestingly, Irene also preferred some challenging tasks such as interpreting and finding the subtle meaning differences between synonyms. She recalled that distinguishing *hear* and *listen*, *end* and *finish* was interesting because she was able to understand the contexts of using each word appropriately. In this step, Irene spent more time reading and interpreting each sentence. She tried her best to interpret all sentences but she used online translator, Papago, for most of the time. Even though she used translator, her overall reaction toward this activity was that it helped her a lot to distinguish the differences of synonym pairs correctly and to develop self-directed learning skill.

In the last session, Irene was required to compare and contrast three words at once – *look, see, watch*. Within 45 minutes of the whole class time, Irene could not finish the worksheet completely. She recalled that the reason was not because it was difficult but because there were many things to write down. Indeed, not many intervals were shown and Irene kept working on for the whole 45 minutes. She finished organizing the usage of word *look* within 10 minutes but then it took longer when she searched the word *see* and *watch*. While reviewing what she has done, she recalled:

> Excerpt 4.9. When I entered the word *watch*, I noticed that words such as TV or movie follows *watch*. However, I could not characterize common usages of the word *see* in a limited time so at this point, I stopped and listened to the hint that teacher gave us. The teacher's hint helped me to complete the worksheet. (Irene's Interview Quote, Session 3)

Irene added her comment that comparing two words seems appropriate for her in 45 minutes. Although she said she exactly understands how to search the differences between synonyms using corpus, completing a task with three words was burdensome. She ended the interview that she might have done herself without a support if she had more time to complete the worksheet.

4.3.2.2. Performance on Using Computer

Irene showed a very positive reaction toward using computer in class. At first, she described the most impressive point of the lesson as writing down what she found using her own computer. Irene recalled it was so interesting because she has never used computer during regular classes before. She emphasized that using computer is very comfortable because she can easily copy and paste what she found to her worksheet. The only thing she commented negatively was difficulty of editing the worksheet. Although she could type English without any problem, she was not used to editing the file neatly so she spent about three minutes editing it before she submitted the worksheet on Padlet.

Both in the second and third week, Irene did not show any problem regarding computer-assisted language learning. In her third interview, she recalled:

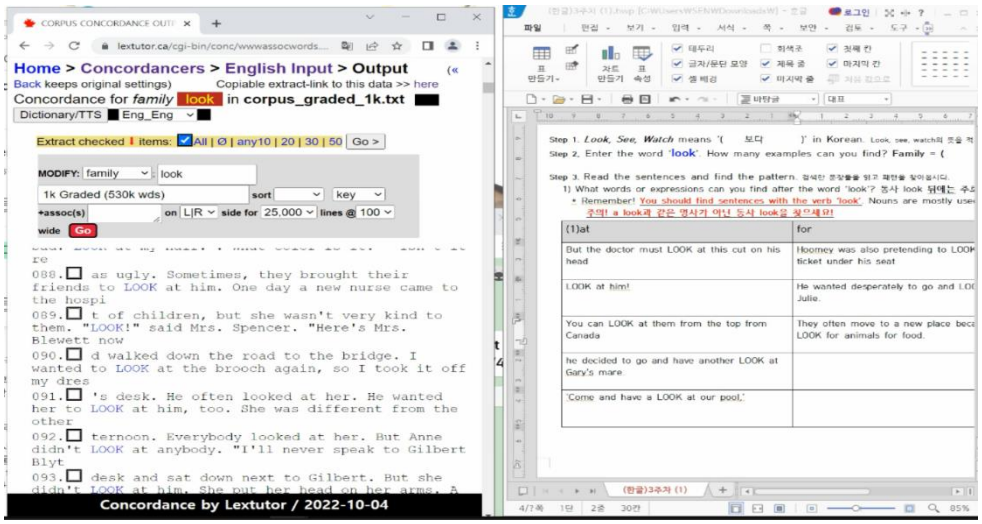
> Excerpt 4.10. As I kept using computer during class, I recognized that I really like using computer when learning English. It is just fun and I can find something that I am curious about right away. I like everything. I can now easily type in and edit the worksheet, too. I hope I can use computer in every class. (Irene's Interview Quote, Session 3)

The interview with Irene and her performance showed that she adapted to using computer very well. Figure 4.3 presents a screenshot from Irene's screen recording. Unlike other students, Irene discovered her own efficient way of

completing the task. As the task required students to search examples from *Lextutor* and type what they found in the worksheet, Irene reduced the size of each window so that she could look at both windows at once. She explained that she could perform better and faster based on this method.

Figure 4.3.

Screenshot of Irene's Task



4.3.2.3. Perception on Computer-assisted DDL

Irene described that her perception on computer-assisted DDL changed over time. She recalled:

Excerpt 4.11. I first worried a lot about completing the task myself using computer. It was because I have never done this before and did not know

what and how to do it. Also, I did not have confidence in studying independently so I doubted myself and the lesson. (Irene's Interview Quote, Session 1)

However, her perspective on computer-assisted DDL has gradually changed. She expressed curiosity toward every process and material of the lesson. For example, she wondered about who developed corpus data and website such as *Lextutor* as it was useful to her. Moreover, she mentioned that she could complete the task with greater confidence in the second and third week, which led to better understanding of synonymous verbs. Irene recalled:

Excerpt 4.12. I think the whole lesson procedure was an effective way of learning synonyms. I did not exactly know the differences between synonym pairs at the pretest but today, the questions were very familiar to me. I fully understand what I have learned and I was surprised that I did not forget the differences. Maybe it is because I searched and figured out myself, not with a teacher's one-way explanation. (Irene's Interview Quote, Session 4)

As Irene mentioned, she described the experiment as a rewarding time for her. She also added some comments about self-directed learning that she experienced throughout the experiment. She pointed out that the possibility of managing her learning process during the lesson is the biggest advantage of

computer-assisted DDL. In detail, she explained:

> Excerpt 4.13. In normal classes, I have to follow the teacher's explanation whether I understand it or not. However, in this lesson, I could slow down when I need more time to think about a point. I think I successfully managed my learning process and developed self-directed learning skills. (Irene's Interview Quote, Session 4)

In addition, she emphasized the usability of computer during class as she could search information through Internet freely and complete the worksheet easily. However, Irene also mentioned some disadvantages of computer-assisted DDL. First, she was not sure whether she figured out the differences correctly. She could only check her work when the teacher gave some feedback or hint during the class. In other cases, she found out some mistakes at the end of the class, which made her to revise and change. Second, she pointed out the fact that she could not always concentrate on the class because of the computer. She was tempted to access other websites during class and she recalled that some students tended to do things that are not related to the class. Lastly, she mentioned some difficulties of operating computer she experienced at the first session. But she pointed out that she could overcome these difficulties next session because she got used to using computer.

4.3.3. Leo: A Learner of Outstanding Development

The last participant, Leo, is a male participant. He started learning English at the age of nine, in the elementary school. He expressed himself as a student who does not like to study English. He said at first, he was interested in learning English, but he started to get left behind in English classes. In the end, it became difficult for him to understand the texts and contents taught in school. Although he has a low proficiency of English, he shows a will to learn English during the whole experiment and he became a learner who shows an outstanding growth throughout the session. A detailed behavior of Leo based on the interview and screen recording is described in the following sections.

4.3.3.1. Performance on Tasks

In the first interview, Leo was requested to express the difficulty of the task. Leo answered that entering the words in *Lextutor*, figuring out the number of examples and writing down the sentences in the worksheet were not difficult but organizing similarities and differences of the verb pairs was difficult for him. Leo recalled:

> Excerpt 4.14. At first, I could not understand the task so I listened to the teacher's explanation and followed the direction. For example, I entered the word in the website first, tried to read and interpret the sentences there

but could not characterize the usage myself. So I waited until the teacher gave us a hint. (Leo's Interview Quote, Session 1)

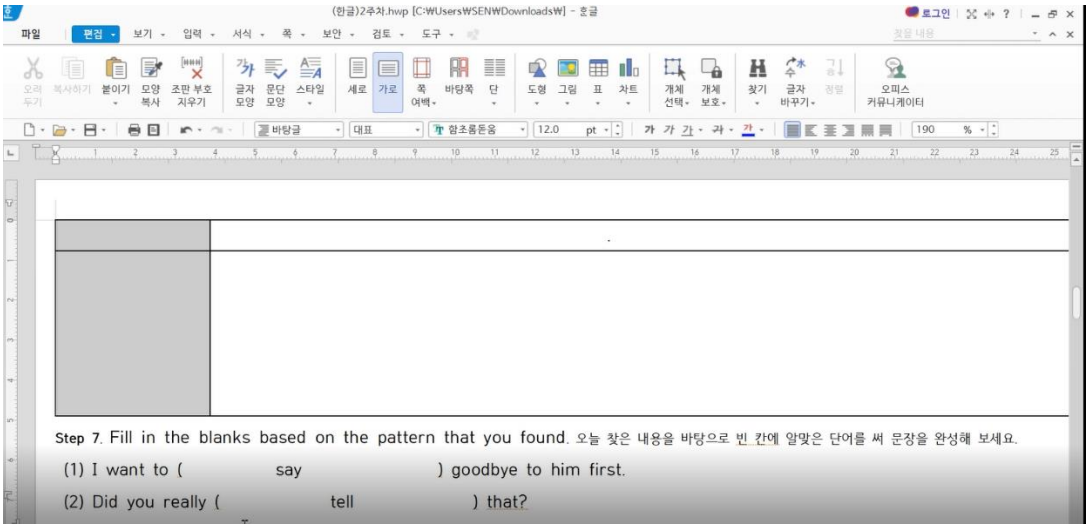
The screen recording of Leo's first task proved what he described. Leo's performance on screen tended to stop for a minute frequently during the 45-minute session. These intervals were the point where Leo was waiting for the teacher's direction. For example, in case of *speak*, Leo could characterize the use of language after *speak* when the teacher highlighted the words such as English and Chinese. In addition, it took a lot of time for Leo to complete all eight steps provided by the researcher so he could not finish the first worksheet in 45 minutes. He explained that he did not have enough time to make his own sentences as he was not familiar with a self-directed learning.

In the second week, Leo showed a development. He explained that he tried his best to complete the task on his own. When figuring out the differences between *say* and *tell*, Leo easily found that *tell* goes with person object. He recalled that as he entered the word *tell*, he could immediately observe expressions such as me, him, or her. However, it was difficult for him to organize the difference between *say* and *tell* in one word – message-focused and addressee-focused – so Leo waited until the teacher gave some hints. He recalled that after a short hint from the teacher, he was able to complete the worksheet himself. But still, as Figure 4.4 shows, Leo's worksheet was simple and there were some mistakes, too.

Leo failed to fill in the blanks based on the differences that he found during the lesson at first. He then revised his answer after the researcher's explanation.

Figure 4.4.

Screenshot from Leo's Worksheet



The researcher then asked Leo about the task of translating concordance lines from *Lextutor*. Leo answered that he could not translate the sentences himself so he used online translator every time. He showed a positive reaction toward translation as he could understand the meaning of English sentences perfectly. He added his comment that by reading each sentences carefully, he could notice the structure of English sentences better than before. Leo said:

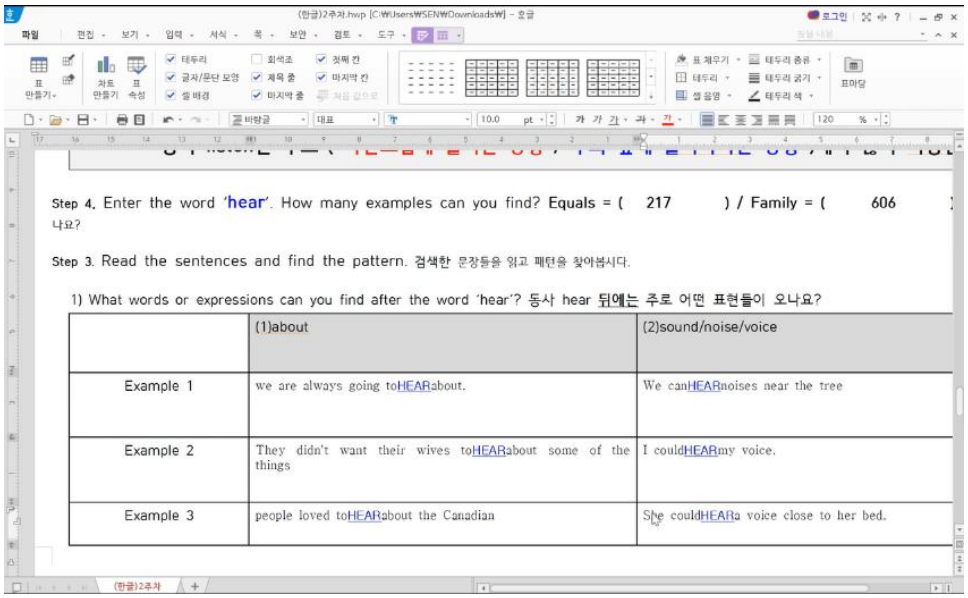
> Excerpt 4.15. I tried my best to complete the task myself. Compared to last week, I clearly understood each step of the task and got used to it. Although I was slow compared to other students in my opinion, I was sure

that I can explain the differences between the synonym pairs when my friend asks me about it and I learned a lot through the lesson. (Leo's Interview Quote, Session 2)

As Leo's words shows, he became confidence himself from the second week. Figure 4.5 shows Leo's performance of completing the task. Although what he found was not a primary usage that the researcher intended, he wrote down the expressions that he observed based on concordance lines on his own.

Figure 4.5.

Screenshot of Leo's Task



Step 4. Enter the word 'hear'. How many examples can you find? Equals = (217) / Family = (606)
나요?

Step 3. Read the sentences and find the pattern. 검색한 문장들을 읽고 패턴을 찾아봅시다.

1) What words or expressions can you find after the word 'hear'? 동사 hear 뒤에는 주로 어떤 표현들이 오나요?

	(1)about	(2)sound/noise/voice
Example 1	we are always going toHEARabout.	We canHEARnoises near the tree
Example 2	They didn't want their wives toHEARabout some of the things	I couldHEARmy voice.
Example 3	people loved toHEARabout the Canadian	She couldHEARa voice close to her bed.

In the last session, Leo was required to compare three verbs – *look, see, watch* – at once. Leo easily found the expression *look at* in about 3 minutes by

himself. Then, he chose five example sentences and interpreted them for another 7 minutes. For the next word, *see*, as the participants were required to only translate example sentences and find out in what context the word *see* is used, Leo also started translation. However, he started to hesitate. His screen stopped for more than 2 minutes. Leo later recalled that he waited the teacher's hint because he could not think of where and how the word *see* is used. For the last word, *watch*, Leo could not fill all the blanks in the worksheet. In this step, the researcher purposefully read some expressions found in *Lextutor*, such as 'watch TV', or 'watch movie'. Based on the hints from the teacher, he tried to find some similar expressions in concordance lines and figured out the primary usage of the word. Leo expressed that although he submitted the worksheet, he was not sure that he fully understood the contents because he was out of time in the end. Leo wanted to review it during the interview so he searched again and internalize the differences with some questions to the researcher. In the end, Leo replied that although hints and support from the teacher was necessary for him, he learned a lot through the lesson and self-directed tasks.

4.3.3.2. Performance on Using Computer

Leo experienced some difficulties of using computer in the first session. He responded that he has never used computer for the purpose of learning and also

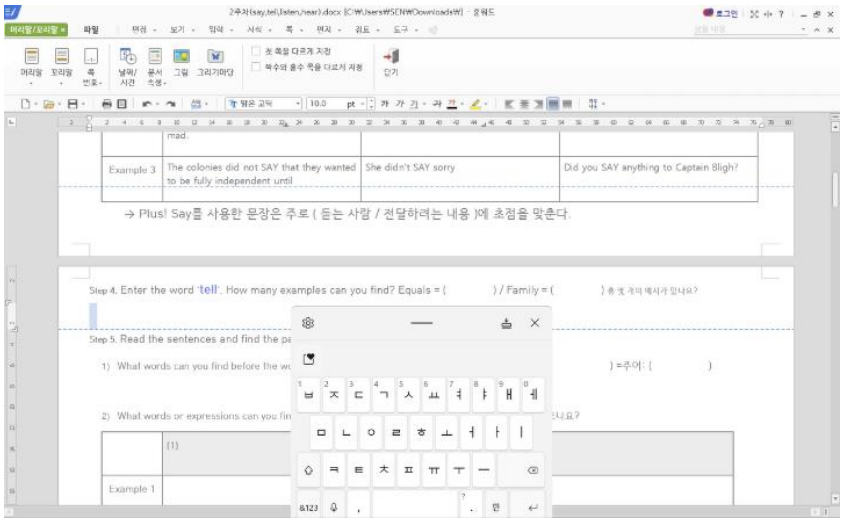
he was not used to operating the tablet PC provided by school. As he could not manage computer well, it took a lot of time to download and save the worksheet in his computer. He recalled:

> Excerpt 4.16. As I did not have much experience of working on a document on the computer, I think I need a practice session regarding some basic skills of operating computer. I asked the teacher and friends to solve my problem today, but my computer did not work well. There were many errors in my computer so I couldn't follow up the lesson properly because of those errors. (Leo's Interview Quote, Session 1)

In the second and third week, Leo reported that his computer worked well. He clearly understood how to download, save and submit files on computer. Moreover, he discovered his own way of converting windows easily so he did not show any problems of using computer. The only inconvenience he expressed was typing and editing using Microsoft Word. He recalled that at some point, the tables in the worksheet suddenly disappeared or the screen showed certain function that he did not intend to. Figure 4.6 shows the capture of Leo's worksheet where Leo stopped about a minute because of an error while typing the worksheet. However, Leo highlighted the usefulness of computer while completing the tasks as those errors were minor things and easily resolved.

Figure 4.6.

Screenshot of Leo's Worksheet



4.3.3.3. Perception on Computer-assisted DDL

Although Leo experienced some difficulties while completing the task throughout the whole session, his perception on computer-assisted DDL was very positive. At the last interview, Leo described about the posttest. He reported that compared to the pretest, which he did not know anything about the differences between synonyms, he understood all the questions and tried to remember what he learned during the experiment. He was satisfied with the result, mentioning that:

> Excerpt 4.17. I was surprised at myself because when I read the sentences in the posttest, I could recall what I found and learned during that session. That was why I did not guess any question in the posttest. (Leo's Interview Quote, Session 4)

Leo was also requested to explain the positive and negative aspects of data-driven learning, an example of self-directed learning. He replied that data-driven learning helped him to develop autonomy. He has only followed the teacher's direction before, wrote what the teacher explained during classes so he did not feel he was learning. However, the whole lesson procedure encouraged him to solve the problem on his own, even though he got wrong sometimes. The negative aspect he pointed out was related to the task difficulty. He said it was challenging for him to complete the task every class and he could not finish it without the teacher's support. He worried that the students who are not proficient in English like him would give up as they could not understand the task and the lesson.

In case of CALL, Leo showed positive responses. He mentioned that the biggest advantage of CALL is the possibility of searching what they do not know through Internet, which was not allowed in normal classes. He hoped to use their own educational tablet PC every class so that he can search the vocabulary he does not understand or use online translator to translate some complex sentences by himself. Leo also described that by using computer, English class become more interesting to him and he could concentrate on class better. The only disadvantage of CALL was some errors occurred in the computer. However, Leo explained that such technical problems were not significant as he adapted to using computer

slowly. In the end, he suggested a practice session to become proficient at functions of computer, which would be helpful to many students.

4.3.4. Summary of the Descriptive Accounts

Through an examination of individual interviews and screen recordings, an evaluation was conducted on the performances of three participants: Hanna, Irene, and Leo. A prominent observation emerged wherein all participants acknowledged the efficacy of computer-assisted DDL in discerning synonymous verbs. Despite their initial encounter with corpora, all three participants concurred on the substantial pedagogical value of corpora as a language learning tool. Furthermore, a shared inclination towards incorporating computers into English lessons was discernible among them.

Nonetheless, discernible variations in behavior and responses were exhibited by each participant throughout the session. These discrepancies stemmed from disparities in the participants' levels of English proficiency and preferred learning styles. Proficient learners demonstrated a propensity for adeptly engaging in self-directed learning, consistently completing tasks even in the absence of external assistance. Learners with intermediate proficiency encountered some challenges in autonomously deducing linguistic rules; however, they successfully differentiated synonymous verbs with the guidance and support

provided by the teacher. Conversely, the self-directed tasks proved arduous for learners with low proficiency, necessitating substantial hints and support from the teacher. Despite the inherent difficulties encountered, learners with low proficiency attested to the efficacy of computer-assisted DDL as an effective pedagogical approach.

CHAPTER 5. DISCUSSION

This chapter discusses a detailed analysis of key research findings presented in Chapter 4, in relation to previous studies. Section 5.1 presents the effectiveness of implementing computer-assisted DDL on pre-tertiary learners, along with some additional suggestions. In Section 5.2, the discussion focuses on important ideas from the Sociocultural Theory (SCT) framework when implementing computer-assisted DDL.

5.1. Implementation of Computer-assisted DDL on Pre-tertiary Learners

The quantitative results of the present study indicate that computer-assisted DDL was proven to be an effective methodology, particularly for pre-tertiary learners.

First, the implementation of computer-assisted DDL led to improvements in pre-tertiary learners' ability to differentiate synonymous verbs. Previous research studies have consistently emphasized the importance of paying special attention to synonymy in language learning contexts (Jung et al., 2007; Liu & Zhong, 2016; Yevchuk, 2022). However, there is a dearth of experimental studies utilizing corpora as an educational tool for learning synonyms (Yeh et al., 2007; Kim, 2020). The

present study successfully employed corpora as an educational tool for learning synonyms and verified that Korean middle school EFL learners can effectively distinguish synonymous verbs through the implementation of DDL.

Furthermore, based on Gabrielatos' (2005) categorization of DDL, the utilization of the hard version of DDL, which involves direct access to raw corpus data using computer facilities, was found to be effective for pre-tertiary learners across different proficiency levels in the present study. These findings support the argument put forth by Boulton (2010) that DDL can be successfully applied to learners at the beginner level of language proficiency. Therefore, it can be concluded that the hard version of DDL, combining both DDL and CALL, serves as an effective teaching methodology for distinguishing English synonyms, as evidenced by the results of this study.

The results of the questionnaire, which included 5-point Likert scale questions and open-ended questions, indicate that learners perceive computer-assisted DDL as a helpful and interesting methodology for learning English. The learners highlighted the usefulness of corpora as a means to access authentic examples and expressed interest in CALL as they could freely use computers during class. These reactions support the advantages of DDL and CALL mentioned in previous studies. One of the recognized advantages of DDL is providing authenticity to language learners (Johns, 1997; Chambers, 2010; Gliquin & Granger, 2010).

Middle school EFL learners in this study tended to agree with this advantage, resulting in high mean scores for the question related to the authenticity of corpus data. Moreover, CALL fostered a learner-friendly learning environment, which resulted in positive reactions from the learners.

However, some learners encountered difficulties and expressed concerns regarding computer-assisted DDL. These challenges were attributed to the complexity of using computers and analyzing corpus data for some learners. The results support the limitations identified in previous studies, where some learners exhibited negative reactions towards DDL due to their lack of familiarity with using computers for learning purposes (Liu et al., 2002). The importance of training sessions focused on both DDL and CALL should be emphasized in order to address the difficulties encountered by pre-tertiary learners. As argued in previous studies, initial training sessions should be conducted prior to actual DDL sessions, covering the methodology of DDL and the analysis and interpretation of corpus data (Liontou, 2020). Furthermore, considering that pre-tertiary learners do not possess automatic skills for effectively utilizing digital tools and materials for language learning, learner training on utilizing digital tools and materials should be prioritized (Hubbard & Romeo, 2012).

In fact, the present study provided a training session with a detailed lesson procedure for DDL, based on the three-step procedure for concordance-based

learning suggested by Johns (1991) and the framework of corpus-based studies proposed by Flowerdew (2009). This session assisted participants in word searching and data interpretation within the corpus. Contrary to this, the emphasis of training session was not significantly placed on CALL and the various methods of computer operation during the lesson. Consequently, some participants, including the participant with low proficiency, expressed difficulties in completing the worksheet using their own educational devices. This demonstrates the need for a comprehensive training session when implementing computer-assisted DDL for pre-tertiary learners.

The final aspect to be taken into account during the implementation of computer-assisted DDL pertains to the learning style of the learners. While the behavior and responses of the learners were primarily distinguished by their level of English proficiency, their learning style appeared to have an impact on their task performance. Within the sample of three participants, both high and low proficiency learners exhibited traits of active learners who actively engaged in class by posing questions and expressing their opinions. Conversely, the intermediate learner demonstrated characteristics of a reflective learner. She required ample time to fully comprehend the corpus data, and expressed difficulty in seeking assistance from the instructor. This finding highlights the imperative of considering the individual learning styles of learners when introducing new type of teaching and learning

methodologies. The efficacy of a specific learning methodology can vary among learners based on their individual learning styles. Consequently, instructors should consider learners' learning styles when incorporating DDL lessons into their teaching practices.

5.2. Computer-assisted DDL and the SCT Paradigm

Despite the demonstrated effectiveness of computer-assisted DDL as a methodology for teaching synonymous verbs to Korean EFL middle school students, the qualitative findings of this study underscore the significance of analyzing DDL within the framework of the SCT paradigm. As expounded upon in Section 2.2.1.1, the SCT paradigm places particular emphasis on the concept of 'scaffolding.' Previous studies have consistently emphasized the necessity of scaffolding during DDL, as it minimizes cognitive demands on learners (Cobb & Boulton, 2015; Flowerdew, 2009, 2015). Within the SCT paradigm of DDL, the process of mediation and scaffolding by peers or an instructor is essential for learners to become self-regulated. In the present study, computer-assisted DDL was implemented with a target form selected by the instructor, and interventions throughout the sessions aided learners in successfully acquiring the target form.

Specifically, the descriptive analysis of the three participants revealed the essence of DDL, wherein students assume the role of 'language detectives' (Johns,

1997), independently uncovering linguistic facts and rules based on authentic examples. However, the analysis also highlighted the necessity of mediation for intermediate and low proficiency learners.

Firstly, the high proficiency learner swiftly grasped the process of distinguishing synonymous verbs using corpus data and consistently completed the computer-assisted DDL tasks in each session. By the second and third week, she demonstrated the ability to discern synonymous verbs autonomously, drawing from the authentic examples provided by the corpus data, and even relished some challenging tasks. In fact, she required minimal scaffolding or intervention from the instructor. Her overall behavior and responses attest to the suitability of computer-assisted DDL as a pedagogical approach for her, fostering her development as an autonomous learner.

Secondly, the intermediate proficiency learner initially encountered difficulties in adapting to the process of distinguishing synonymous verbs. For instance, she struggled to complete tasks on time when faced with challenging target words and frequently relied on the instructor's scaffolding and hints. In detail, the instructor placed emphasis on certain concordance lines, which helped participants to distinguish the synonym pairs. Nevertheless, she exerted considerable effort in utilizing the corpus data and comprehending the distinctions between the words. Despite finding some tasks burdensome, she expressed enjoyment during the

sessions and particularly favored using the computer. Additionally, she emphasized the significance of the support provided by the instructor, which greatly facilitated her understanding of the differences between synonyms. In comparison to the high proficiency learner, the intermediate proficiency learner necessitated more scaffolding throughout the sessions.

Lastly, the low proficiency learner heavily depended on the instructor's scaffolding, explanations, and hints throughout the sessions. The low proficiency learner voiced difficulties in differentiating synonymous verbs based on the corpus data, as comprehension of English sentences posed a challenge. He often failed to complete tasks within the allotted time and frequently relied on the instructor's support. Despite perceiving computer-assisted DDL as burdensome, including the utilization of computers, the learner demonstrated progress over time. For instance, in the second and third sessions, he managed to independently complete the basic steps of the task. Furthermore, he recalled being able to explain the differences between the synonym pairs to others, indicating successful comprehension of the target learning items. At the conclusion of the sessions, he exhibited significant development, validating the effectiveness of computer-assisted DDL in distinguishing synonymous verbs.

The findings of this study highlight the connection between the observed processes and the SCT paradigm of DDL, an area of exploration that has garnered

attention from researchers (O'Keeffe, 2021; Pérez-Paredes, 2022). While the primary objective of DDL is to facilitate students' independent discovery of target forms or expressions, the presence of mediation during DDL lessons is crucial for the successful implementation of DDL across all proficiency levels of language learners. Particularly, as evidenced by the experiences of low and intermediate proficiency students in this study, the provision of instructor scaffolding enabled them to gain control over the entire learning process. It is worth noting that the role of intervention or mediation extends beyond merely providing hints. It involves assisting learners in becoming self-directed learners.

In conclusion, computer-assisted DDL, a form of DDL as a hard version, was found to be effective for Korean middle school EFL learners. The participants developed their ability to distinguish synonyms following the sessions and demonstrated a positive perception of implementing computer-assisted DDL in the classroom. However, advanced learners exhibited a higher aptitude for adapting to computer-assisted DDL tasks, as they could comprehend the tasks and authentic examples without requiring scaffolding. On the other hand, learners with relatively low English proficiency relied on hints provided during each session as scaffolding. Based on the theoretical framework of the SCT paradigm, these overall processes of computer-assisted DDL including mediation for certain learners ultimately foster the development of 'language detectives' (Johns, 1997), thereby ensuring the

successful implementation of computer-assisted DDL for pre-tertiary learners of English.

CHAPTER 6. CONCLUSION

The final chapter concludes the present study by summarizing the major findings, implications, and limitations. In Section 6.1, major findings and their implications of the study are discussed in relation to the three research questions. Then, Section 6.2 presents the limitations of the study and provides some suggestions for future studies.

6.1. Major Findings and Implications

The major goal of the present study was to investigate the effects of computer-assisted DDL on distinguishing synonyms, in the case of Korean EFL middle school students. In detail, this study examined both the cognitive and affective domain by answering the following three research questions: 1) Do Korean EFL middle school students improve their comprehension ability to distinguish synonymous verbs through computer-assisted DDL? 2) How do learners perceive computer-assisted DDL? 3) What changes do learners experience in their cognitive and affective domains through computer-assisted DDL?

Regarding the first research question, the pretest and the posttest results from the experiment showed that computer-assisted DDL was effective for Korean EFL middle school on distinguishing synonymous verbs as their average posttest

score was higher compared to the pretest score. This result indicates that DDL is possible with all learners (Boulton & Cobb, 2017). In addition, as Boulton and Cobb (2017) argued that there is an evolution towards practicing hard version of DDL recently, the result from the experiment supports the possibility of implementing an actual digital turn in DDL.

For the second question, the results of the questionnaire examined that Korean EFL middle school students generally tended to show positive perception toward computer-assisted DDL. Many students responded that reading authentic examples helped them to expand their vocabulary knowledge, which is one of the advantages of DDL pointed out by many previous researchers (e.g., Johns, 1997; Chambers, 2010; Gliquin & Granger, 2010). Moreover, most of the learners were encouraged to become active participants in learning, expressing that DDL activities were interesting compared to ordinary vocabulary classes. Although a small number of participants expressed difficulties using computer throughout the whole class and discovering the differences between synonym pairs, the result of this questionnaire supports the findings of previous studies that DDL and CALL guarantees language learners' positive attitudes (e.g., Crosthwaite & Stell, 2020; Felix, 2005a; Pérez-Paredes, 2022).

Lastly, for the third research question, individual learners showed different behavior and responses throughout the interview, but the descriptive accounts

indicate that despite language proficiency level, all learners experienced their growth and development. In detail, high proficiency learner successfully completed computer-assisted DDL task every session and enjoyed interpreting corpus data. In contrast, intermediate and low proficiency learner required scaffolding to understand and complete the task, experiencing some difficulties when the differences between synonym pairs were not obvious. As many previous studies argued that applying DDL to pre-tertiary learners who have low proficiency level is burdensome (Yoon & Hirvela, 2004; Chambers, 2005; Boulton, 2009), the result of the present study also showed the necessity of providing further support to those students (Chang, 2012).

The current study is meaningful in three aspects. First, it enlightened the possibility of implementing DDL to pre-tertiary learners. In the language education field, almost all DDL studies were conducted with adult learners with high language proficiency level (Boulton & Cobb, 2017; Pérez-Paredes, 2022). This tendency prevented the development of DDL studies in secondary schools. However, this study proved that DDL is effective for beginning level of language learners (e.g., Boontam & Phoocharoensil, 2018; Crosthwaite & Steeples, 2022) when the tasks and corpus tools are suitable to those learners. Second, the study of computer-assisted DDL has never been conducted in Korean secondary language education context. Although CALL is comparably popular and common way of teaching and

learning English, the studies of DDL in CALL researches are not yet a primary area of practice (Pérez-Paredes, 2022). The current study tended to overcome the sample bias in DDL studies, applying hard version of DDL based on CALL to pre-tertiary learners in Korea. Lastly, this study involved qualitative analysis. Almost all previous researches of DDL focused on examining the effects of corpus-based learning by conducting quantitative experiment, involving large samples (Pérez-Paredes, 2022). However, examining the individual learner's learning process during DDL is meaningful as the researcher can find out how the learner completes the task, changes, and feels by experiencing DDL in detail. Since this study is based on both quantitative and qualitative methodology, the findings provide better understandings of how learners benefit from DDL.

The findings from this study also have an implication for language teaching and learning in Korean pre-tertiary school EFL contexts. As previously mentioned in Section 1.1, the educational environment of Korea started to change. Middle school 1st grade students in Seoul received individual computer device, so that they can freely use computer and Internet during classes. As a result, one of the biggest limitations of implementing hard version of DDL, the problem of logistics, got solved. It becomes possible to implement the hard version of DDL, or computer-assisted DDL in Korean middle school English classes, which is proved to be effective in this study. Therefore, the findings in the present study are expected to

shed light on applying DDL to pre-tertiary learners regardless of their English proficiency level. Instead of explaining the differences between synonym pairs directly to the students, providing the opportunities to discover the differences themselves based on corpus data should be encouraged. Although DDL may require more time compared to traditional teacher-centered instruction, it fosters self-directed and autonomous learners. Through this process, learners take on an active role in their own learning, aligning with the ultimate goals of DDL, namely, becoming 'language detectives' (Johns, 1997) and developing learner agency.

6.2. Limitations and Suggestions for Future Research

Since the current study is not without some limitations, some suggestions for future research are discussed. First, because of the small sample size (N = 29) and regional limitation, the results of this study were insufficient to generalize. For the future research, it is necessary to conduct the experiment with a greater number of participants with diverse backgrounds. Second, this study investigated only the growth between the pretest and posttest. To examine the effects of computer-assisted DDL on distinguishing synonymous verbs more accurately, a delayed posttest should be conducted few weeks after the posttest. The differences between the posttest and delayed posttest should be examined to prove the long-term effects of computer-assisted DDL. Third, the participants' learning style was not considered

in the present study. As one of the keywords of DDL is variety (Gliquin & Granger, 2010), it is important to admit the learners' different learning styles and preferences. The future studies considering the learning style should be implemented. Lastly, this study targeted only five pairs of synonymous verbs. Therefore, it is necessary to conduct future studies dealing with other types of vocabulary or target items to generalize the effects of computer-assisted DDL.

Despite these limitations, the present study contributes to the existing literature on DDL by demonstrating the positive impact of employing a challenging version of DDL, utilizing individual computers, on the proficiency of Korean middle school learners in distinguishing synonymous verbs. Given the increasing number of studies that have begun to explore DDL for pre-tertiary learners and to implement hard version of DDL, further research and recommendations pertaining to DDL, applicable to learners of all proficiency levels, would be beneficial for language educators in the future.

REFERENCES

Alanazi, M. (2017). The Comprehension of Synonyms by Saudi EFL Learners: Acquisition and Pedagogical Implications. *International Journal of Applied Linguistics & English Literature, 6*(3), 58.

Allum, P. (2002). CALL and the classroom: the case for comparative research. *ReCALL, 14*(1), 146-166.

Arppe, A., & Järvikivi, J. (2007). Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory, 3*(2), 131-159.

Barrette, C. (2001). Students' preparedness and training for CALL. *CALICO Journal, 19*(1), 5-36.

Bernardini, S. (2001). "Spoilt of Choice": A Learners Explores General Language Corpora. In G. Aston (Ed.), *Learning with Corpora* (pp. 220-249). Athelstan.

Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English preposition learning through data-driven learning (DDL). *3L, Language, Linguistics, Literature, 24*(3), 125-141.

Boulton, A. (2008). DDL: Reaching the Parts Other Teaching Can't Reach? In A. Frankenburg-Garcia (Ed.), *Proceedings of 8th Teaching and Language Corpora Conference*. Lisbon: Associação de Estudos e de Investigação Cientifica do ISLA-Lisboa, p. 38-44.

Boulton, A. (2009). Data-driven Learning: Reasonable Fears and Rational Reassurance. *Indian Journal of Applied Linguistics, 35*(1), 81-106.

Boulton, A. (2010). Language awareness and medium-term benefits of corpus consultation. In A. Gimeno Sanz (Ed.), *New trends in computer-assisted language learning: Working together* (pp. 39-46). Macmillan ELT.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393.

Boulton, A., & Pérez-Paredes, P. (2014). ReCALL special issue: Researching uses of corpora for language teaching and learning Editorial Researching uses of corpora for language teaching and learning. *ReCALL, 26*(2), 121-127.

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

Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning, 22*(2), 153-172.

Bruner, J. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. Jarvella and W. J. J. Levelt (Eds.) *The Child's Conception of Language* (pp. 241-256). Spring-Verlag.

Chambers, A. (2005). Integrating Corpus Consultation in Language Studies. *Language Learning and Technology, 9*(2), 111-125.

Chambers, A. (2010). What is data-driven learning? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 345-358). Routledge.

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

Chambers, A., & Bax, S. (2006). Making CALL work: Towards normalization. *System, 344*, 465-479.

Chambers, A., Farr, F., & O'Riordan, S. (2011). Language teachers with corpora in mind: From starting steps to walking tall. *The Language Learning Journal, 39*(1), 85-104.

Chang, P. (2012). Using a stance corpus to learn about effective authorial stance-taking: A textlinguistic approach. *ReCALL, 242*, 209-236.

Chapelle, C. (2003). *English language learning and technology*. John Benjamins.

Charles, M. (2007). Reconciling Top-down and Bottom-up Approaches to Graduate Writing: Using Corpus to Teach Rhetorical Functions. *Journal of English for Academic Purposes, 6*, 289-302.

Choi, S. (1996). The effectiveness of CD-ROM courseware in a high school setting in Korea. *English Teaching, 51*(4), 177-202.

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*(2), 132-145.

Cobb, T. (1997). *From concord to lexicon: Development and test of a corpus-based lexical tutor* [Unpublished doctoral dissertation]. Concordia University.

Cobb, T. (1999). Applying constructivism: A test for the learner as scientists. *Educational Technology Research & Development, 47*(3), 15-31.

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber and R. Reppen (Eds.), *Cambridge handbook of corpus linguistics* (pp. 478-497). Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Crosthwaite, P. (2020) Data-driven learning and young learners. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 1-10). Routledge.

Crosthwaite, P. & Stell, A. (2020). "It helps me get ideas on how to use my words": primary school students' initial reactions to corpus use in a private tutoring setting. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 150-170). Routledge.

Crosthwaite, P. & Steeples, B. (2022). Data-driven learning with younger learners: exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*, DOI:10.1080/09588221.2022.2068615.

Cruse, D. (2000). *Meaning in Language. An Introduction to Semantics and Pragmatics.* Oxford University Press.

Davies, G., Otto, S. E., & Rüschoff, B. (2014). Historical perspectives on CALL. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 19-38). Bloomsbury.

Dirven, R., Goossens, L., Putseys, Y., & Vorlat, E. (1982). *The scene of linguistic action and its perspectivization by speak, talk, say and tell*. John Benjamins Publishing.

Doughty, C., & Long, M. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning and Technology, 7*(3), 50-80.

Edmonds, P., & Hirst, G. (2002). Near-Synonymy and Lexical Choice. *Computational Linguistics - Association for Computational Linguistics, 28*(2), 105-144.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1-24.

Farr, F. & Murray, L. (2016). Language learning and technology. In Farr, F., & Murray, L. (Eds.) *The Routledge Handbook of Language Learning and Technology* (1st ed., pp.1-6). Routledge.

Felix, U. (2005a). What do meta-analyses tell us about CALL effectiveness? *ReCALL, 17*(2), 269-288.

Felix, U. (2005b) Analyzing Recent CALL Effectiveness Research – Towards a Common Agenda. *Computer Assisted Language Learning, 18*(1–2): 1–32.

Firth, J. (1957). *Papers in Linguistics 1934-1951*. Oxford University Press.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning, 20*(4), 409-442.

Fischer, R. (2012). Diversity in learner usage patterns. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 14-32). Cambridge University Press.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics, 14*(3), 393-417.

Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain shall meet. In A. Leńko-Szymańska and A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15-36). John Benjamins.

Forti, L. (2018). *Developing phraseological competence in Italian L2: A study on the effects of Data-driven Learning* [Unpublished doctoral dissertation]. University of Foreigners of Perugia.

Frankenberg-Gracia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL, 26*(2), 128-146.

Freed, A. F. (1979). *The Semantics of English Aspectual Complementation* (Synthese Language Library, 8.). D. Reidel.

Friginal, E. (2018). *Corpus Linguistics for English Teachers : New Tools, Online Resources, and Classroom Activities.* Routledge.

Gabrielatos, C. (2005). Corpora and Language Teaching: Just a Fling or Wedding Bells? *TESL-EJ, 8*(4), 1-37.

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL, 262*, 225-242.

Glendenning, E. & Howard, R. (2003). Lotus ScreenCam as an aid to investigating student writing. *Computer Assisted Language Learning, 16*(1), 31-46.

Gliquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-370). Routledge.

Gilquin, G. & Granger, S. (2022). Using data-driven learning in language teaching. In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics. Second Edition* (pp. 430-442). Routledge.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. In *Midwest Research to Practice Conf. in Adult, Continuing, and Community Education*. The Ohio State Univ.

Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL, 25*(2), 165-198.

Gries, S. T., & Otani, N. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal, 34*, 121-150.

Hadley, G. (2002). An Introduction to Data-driven Learning, *RELC Journal, 33*(2), 99-124.

Halliday, M. (1966). Lexis as a linguistic level. In C. Bazell, J. Catford, M. Halliday and R. Robins (Eds.), *In Memory of J. R. Firth* (pp. 148-162). Longman.

Han, Z-H., & Finneran, R. (2013). Re-engaging the interface debate: Strong, weak, none, or all? *International Journal of Applied Linguistics, 24*(3), 370-389.

Hanks, J. (2019). From research-as-practice to exploratory practice-as-research in language teaching and beyond. *Language Teaching, 52*(2), 143-187.

Hatch, E., & Brown, C. (1995). *Vocabulary, Semantics, and Language Acquisition*. Cambridge University Press.

Hirata, E. (2020). The development of a multimodal corpus tool for young EFL learners: a case study on the integration of DDL in teacher education. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 88-105). Routledge.

Hong, S.-Y., & Oh, S.-Y. (2008). The effects of corpus-based learning of vocabulary and grammar on Korean high-school students. *English Language Teaching, 20*(1), 261-283.

Huang, L. (2011). Corpus-aided language learning. *ELT Journal, 65*(4), 481-484.

Hubbard, P. (2004). Learner training for effective use of CALL. In S. Fotos and C. Browne (Eds.), *New Perspectives on CALL for Second Language Classrooms* (pp. 45-68). Lawrence Erlbaum Associates.

Hubbard, P. (2005). A review of subject characteristics in CALL research. *Computer Assisted Language Learning, 18*(5), 351-368.

Hubbard, P. (2008). Twenty-five years of theory in the CALICO Journal. *CALICO Journal, 25*(3), 387-399.

Huiying, Z. (2013). A Comparative Study of Semantically Similar Verbs: Focus on Stop, End & Finish. *Quarterly Journal of Chinese Studies, 1*(4), 94-101.

Hunston, S. (2002). *Corpora in Applied Linguistics* (1st ed.). Cambridge University Press.

Hunston, S. (2022). *Corpora in Applied Linguistics* (2nd ed.). Cambridge University Press.

Indrarathne, B., Ratajczak, M., & Kormos, J. (2018). Modelling changes in the cognitive processing of grammar in implicit and explicit learning conditions: Insight from an eye-tracking study. *Language Learning, 68*(3), 669-708.

Inkpen, D., & Hirst, G. (2006). Building and Using a Lexical Knowledge Base of Near-Synonym Differences. *Computational Linguistics - Association for Computational Linguistics, 32*(2), 223-262.

Jang, B. (2012). Korean EFL Learners' Attitudes toward the Use of CALL and Their Motivations for Learning English in Korea. *The Jungang Journal of English Language and Literature, 54*(4), 437-460.

Johansson, S. (2009). Some thoughts on corpora and second-language acquisition. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 33-44). John Benjamins.

Johns, T. (1986). Micro-Concord: a language learner's research tool. *System, 14*(2), 151-162.

Johns. T. (1988). Whence and Whither Classroom Concordancing? In T. Bongaerts, P. de Haan, S. Lobbe and H. Wekker (Eds.), *Computer Applications in Language Learning* (pp. 9-33). Foris.

Johns, T. (1991). Should you be persuaded - two examples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1-13). Birmingham University.

Johns, T. (1997). Contexts: the Background, Development and Trialling of Concordance-based CALL Program. In Wichmann et al. (Eds.), *Teaching and Language Corpora* (pp. 100-115). Longman.

Jung, Y. (2009). A Corpus-Based Study on the Use of Adverb Synonyms 'Entirely, Totally, and Wholly'. *Linguistic Research, 26*(2), 85-101.

Jung, Y., Yum, J., & Jang, S. (2006). A Corpus-Based Study of Korean Middle School Students' Collocation Ability of English Synonyms. *The New Korean Journal of English Language and Literature*, *48*(4), 247-272.

Jung, Y., Sur, R., & Kim, E. (2007). A Corpus-Based Study of Korean College Students' Collocation Ability of English Synonym Verbs. *The New Korean Journal of English Language and Literature, 49*(2), 175-194.

Kang, N. H. (2019). Effectiveness of collocation learning through a web-concordancer in a Korean EFL university course. *Multimedia-Assisted Language Learning, 22*(1), 41-84.

Kennedy, C., & Miceli, T. (2001). An Evaluation of Intermediate Students' Approaches to Corpus Investigation. *Language Learning and Technology, 5*(3), 77-90.

Kim, H. (2002). Teachers as a barrier to technology-integrated language teaching. *English Teaching, 57*(2), 35-64.

Kim, M. (2020). A qualitative analysis of EFL learns' discrimination of near-synonyms in a data-driven learning task. *English Teaching, 75*(3), 25-47.

Kim, S., & Kim, L. (2021). Happiness education and the Free Year Program in South Korea. *Comparative Education, 57*(2), 247-266.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

Kolaitis, M., Mahoney, M, A., Pomann, H., and Hubbard, P. (2006). Training ourselves to train our students for CALL. In P. Hubbard and M. Levy (Eds.), *Teacher Education in CALL* (pp. 317-332). John Benjamins.

Kulik, J. A. (2003). *Effects of using instructional technology in colleges and universities: What controlled evaluation studies say*. SRI International.

Lasagabaster, D., & Sierra, M. J. (2003). Students' evaluation of CALL software programs. *Educational Media International, 40*(3-4), 293-304.

Laufer, B., & Hulstijin, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1-26.

Lee, D. J. (2013). Data- driven learning (DDL) approach: Is it applicable to Korean EFL classrooms? *Foreign Languages Education, 20*(3), 45- 86.

Lee, H. J, & Lee, E. J. (2010). The effects of corpus-based vocabulary tasks on high school students' English vocabulary learning and attitude. *English Language & Literature Teaching, 16*(4), 239-265.

Lee, H., Warschauer, M., & Lee, J. (2019). The Effects of Corpus Use on Second Language Vocabulary Learning: A Multilevel Meta-Analysis. *Applied Linguistics, 40*(5), 721-753.

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System (Linköping), 81*, 14-25.

Lee, S. (2011). Semantic prosody in bilingual dictionaries and EFL learners' sentence writings. *English Teaching, 66*(2), 253-272.

Leech, G. (1997). Teaching and Language Corpora: A Convergence. In A. Wichmann, S. Fligelstone, T. McEnery and G. Knowles (Eds.) Teaching and Language Corpora (pp. 1-23). Longman.

Leńko-Szymańska, A. (2017). Training teachers in data driven learning: Tackling the challenge. *Language Learning & Technology, 21*(3), 217-241.

Levy, M. (1997). *Computer Assisted Language Learning: Context and Conceptualization*. Clarendon Press.

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer assisted language learning*. Lawrence Erlbaum.

Li, P., & Lan, Y-J. (2022). Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain. *Bilingualism: Language and Cognition 25*, 361–378.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology, 21*(3), 153-171.

Lim, M., & Aryadoust, V. (2021). A scientometric review of research trends in computer-assisted language learning (1977-2020). *Computer Assisted Language Learning, 35*(9), 2675-2700.

Lin, M., & Lee, J. (2019). Pedagogical suitability of data-driven learning in EFL grammar classes: A case study of Taiwanese students. *Language Teaching Research: LTR, 23*(5), 541-561.

Liontou, T. (2020). The effect of data-driven learning activities on young EFL learners' processing of English idioms. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 208-227). Routledge.

Liu, M., Moore, Z., Graham, L., & Lee, S. (2002). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990-2000. *Journal of Research on Technology in Education, 34*(3), 250-273.

Liu, D. (2010). Is it a chief, main, major, primary, or principal concern? A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics, 15*, 56-87.

Liu, D. (2013). Salience and construal in the use of synonymy: A study of two sets of near-synonymous nouns. *Cognitive Linguistics, 24*(1), 67-113.

Liu, D., & Espino, M. (2012). Actually, genuinely, really and truly: A corpus-based behavioral profile study of near-synonymous adverbs. *International Journal of Corpus Linguistics, 43*, 201-226.

Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal, 93*, 61-78.

Liu, D., & Zhong, S. (2016). L2 vs. L1 use of synonymy: An empirical study of synonym use/acquisition. *Applied Linguistics, 37*(2), 239-261.

Mair, C. (2002). Empowering Non-native speakers: The Hidden Surplus Value of Corpora in Continental English Departments. In B. Kettemann and G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 119-130). Rodopi.

Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *The Modern Language Journal, 68*(2), 130-137.

Mauranen, A. (2004). Spoken Corpus for an Ordinary Learner. In J. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 89-105). John Benjamins.

McCarthy, M. (2008). Assesing and Interpreting Corpus Information in the Teacher Education Context. *Language Teaching, 41*(4), 563-574.

McGroarty, M. (1998). Constructive and constructivist challenges for applied linguistics. *Language Learning, 48*, 591-622.

Meunier, F. (2019). A case for constructive alignment in DDL. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 13-30). Routledge.

Ministry of Education. (2013). *Management Plan for Piloting FSP in Middle Schools to Realize Happiness Education by Supporting Students' Dreams and Talents [in Korean]*. Ministry of Education.

Ministry of Education. (2021). *Main Points of Introduction of 2022 Revised Curriculum (Draft Proposal) [in Korean]*. Ministry of Education.

Mizumoto, A., & Chujo, K. (2015). A Meta-analysis of Data-driven Learning Approach in the Japanese EFL Classroom. *English Corpus Studies, 22*, 1-18.

Moon, R. (2010). What can a corpus tell us about lexis? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 197-211). Routledge.

Moon, S., & Oh, S.-Y. (2018). Unlearning overgenerated be through data-driven learning in the secondary EFL classroom. *ReCALL, 30*(1), 48-67.

Morley, J., & Partington, A. (2009). A few frequently asked questions about semantic—or evaluative—prosody. *International Journal of Corpus Linguistics, 14*(2), 139-158.

Mukherjee, J. (2006). Corpus Linguistics and Language Pedagogy: The State of the Art – and Beyond. In S. Braun, K. Kohn and J. Mukherjee (Eds.), *Corpus Technology and Language Pedagogy* (pp. 5-24). Peter Lang.

Mutlu, A., & Eroz-Tuga, B. (2013). The Role of Computer-Assisted Language Learning (CALL) in Promoting Learner Autonomy. *Eurasian Journal of Educational Research, 13*(51), 107-122.

Nim Park, C., & Son, J. B. (2009). Implementing computer-assisted language learning in the EFL classroom: Teachers' perceptions and perspectives. *International Journal of Pedagogies and Learning, 5*(2), 80-101.

O'Bryan, A. (2008). Providing pedagogical learning training in CALL: impact on student use of language-learning strategies and glosses. *CALICO Journal, 26*(1), 142-159.

O'Keeffe, A., McCarthy, M. J., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

O'Keeffe, A. (2021). Data-driven learning – a call for a broader research gaze. *Language Teaching, 54*, 259-272.

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.

Papaioannou, V., Mattheoudakis, M., & Agathopoulou, E. (2020). Data-driven learning in a Greek secondary education setting. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 187-207). Routledge.

Park, E. K. (2011). *Effects of Collaborative Corpus-Based Learning of English Vocabulary by Korean Middle School Students* [Unpublished master's thesis]. Seoul National University.

Park, T. (2011). A Look at Korean EFL Students` Knowledge of the Shades of Meaning in Synonyms. *The Journal of Linguistic Science, 56*, 75-94.

Partington, A. (1998). *Patterns and meanings.* John Benjamins.

Pederson, K. M. (1987) Research on CALL. In: Smith, W. F. (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 99-132). National Textbook Company.

Pérez-Paredes, P. (2010). Corpus Linguistics and language education in perspective: Appropriation and the possibilities scenario. In T. Harris and M. Moreno Jaén (Eds.), *Corpus linguistics in language teaching* (pp. 53-73). Peter Lang.

Pérez-Paredes, P. (2020). The pedagogic advantage of teenage corpora for secondary school learners. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 67-87). Routledge.

Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011-2015. *Computer Assisted Language Learning, 35*(1-2), 36-61.

Pérez-Paredes, P., Ordoñana Guillamón, C., & Aguado Jiménez, P. (2018). Language teachers' perceptions on the use of OER language processing technologies in MALL. *Computer Assisted Language Learning, 31*(5-6), 522-545.

Pérez-Paredes, P., Ordoñana Guillamón, C., Van de Vyver, J., Meurice, A., Aguado Jiménez, P., Conole, G., & Sánchez Hernández, P. (2019). Mobile data-driven language learning: Affordances and learners' perception. *System (Linköping), 84*, 145-159.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912.

Prensky, M. (2011). Digital wisdom and homo sapiens digital. In M. Thomas (Ed.), *Deconstructing digital natives* (pp.25-35). Routledge.

Pujolà, J.-T. (2002). CALLing for help: researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL, 14*(2), 235-262.

Reinders, H., & Darasawang, P. (2012). Diversity in learner support. In G. Stockwell (Ed.), *Computer-Assisted Language Learning: Diversity in Research and Practice* (pp. 49-70). Cambridge University Press.

Reinders, H., & White, C. (2010). The theory and practice of technology in materials development and task design. In N. Harwood (Ed.), *Materials in ELT: Theory and practice* (pp. 58-80). Cambridge University Press

Römer, U. (2006). Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik, 54*(2), 121-134.

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System (Linköping), 91*, 1-14.

Schaeffer-Lacroix, E. (2020). Barriers to trainee teachers' corpus use. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 47-64). Routledge.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 2,* 129-158.

Seong, T.-J. (2016). *Comprehension and application of the item response theory* (2nd Ed.) Kyoyookbook.

Seoul Metropolitan Office of Education. (2022). *2022 Main Educational Tasks in Seoul [in Korean]*. Seoul Metropolitan Office of Education.

Sharwood Smith, M. (1981). Consciousness-raising and the second language Learner. *Applied Linguistics, 2,* 159–168.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Interlanguage studies bulletin (Utrecht), 7*(2), 118–132.

Shi, J. (2014). Pedagogic processing of ESL corpora and its effects on ESL learners' medium-term language awareness enhancement. *International Journal of English Linguistics, 4*(4), 63-73.

Shin, H.-J., & Son, J.-B. (2007). EFL teachers' perceptions and perspectives on Internet-assisted language teaching. *CALL-EJ, 8*(2), 14-31.

Sim, J. (2010). A Comparative Study on the Verbs finish and end and their Collocations. *Studies in Modern Grammar, 62*, 227-252.

Sinclair, J. (1966.) Beginning the study of Lexis. In C. Bazell, J. Catford, M. Halliday and R. Robins (Eds.), *In Memory of J. R. Firth* (pp. 412-429). Longman.

Smith, K., & Craig, H. (2013). Enhancing Learner Autonomy through CALL: A New Model in EFL Curriculum Design. *CALICO Journal, 30*(2), 252-278.

Stockwell, G. (2012). *Computer-assisted Language Learning: Diversity in Research and Practice*. Cambridge University Press.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). Continuum.

Swan, M. (2005). *Practical English usage* (4th Ed.). Oxford University Press.

Szudarski, P. (2018). *Corpus Linguistics for Vocabulary: A Guide for Research*. Routledge.

Szudarski, P. (2020). Effects of data-driven learning on enhancing the phraseological knowledge of secondary school learners of L2 English. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 133-149). Routledge.

Talebinezhad, M., & Abarghoui, M. (2013). The Iranian high school students' attitude toward CALL and the use of CALL for EFL receptive skills. *Theory and Practice in Language Studies, 3*(2), 329-337.

Tanaka-Ellis, N. (2010). Factors limiting learners' success in achieving task outcomes in CALL. *Innovation in Language Learning and Teaching, 4*(3), 213-233.

Taylor, J. (2003). Near synonyms as co-extensive categories: 'high' and 'tall' revisited. *Language Sciences (Oxford), 25*(3), 263-284.

Tribble, C. (2008). From Corpus to Classroom: Language Use and Language Teaching. *ELT Journal, 62*, 213-216.

Tyne, H. (2012). Corpus work with ordinary teachers: Data-driven Learning activities. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 136-151). Masaryk University Press.

Varley, S. (2009). I'll just look that up in the concordance: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning, 22*(2), 133-152.

Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations, *ReCALL, 28*, 207-226.

Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology, 21*(3), 1-8.

Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 144–188). M.E. Sharpe.

Wallwork, A. (2018). Look At, See, Watch. In *Top 50 Vocabulary Mistakes: How to Avoid Them* (pp. 119-120). Springer International Publishing.

Warschauer, M. (2002). Networking into academic discourse. *Journal of English for Academic Purposes, 1*(1), 45–58.

Warschauer, M. (2005). Sociocultural perspectives on CALL. In J. Egbert and G. M. Petrie (Eds.), *CALL Research Perspectives* (pp. 41-51). Lawrence Erlbaum Associates.

Whistle, J. (1999). Concordancing with Students Using an "Off-the-Web" Corpus. *ReCALL, 11*(2), 47-80.

Wicher, O. (2020). Data-driven learning in the secondary classroom: a critical evaluation from the perspective of foreign language didactics. In P. Crosthwaite (Ed.), *Data-Driven Learning for the Next Generation* (pp. 31-46). Routledge.

Wongkhan, P., & Thienthong, A. (2021). EFL Learners' Acquisition of Academic Collocation and Synonymy: Does their Academic Experience Matter? *RELC Journal, 52*(3), 523-538.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103-129.

Yeh, Y., Liou, H., & Li, Y. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning, 20*(2), 131-152.

Yevchuk, A. (2022). An Empirical Study of Near-synonym Choice: A Comparison of Advanced EFL Learners to L1 English Speakers. *Taikomoji kalbotyra, 17*, 79-94.

Yi, J., & Cha, K. W. (2016). Effectiveness of CALL Software for EFL Learners' Improvement of Speaking Proficiency. *Multimedia-Assisted Language Learning, 19*(3), 61-82.

Yoon, H., & Hirvela, A. (2004). ESL Student Attitudes Towards Corpus Use in L2 Writing. *Journal of Second Language Writing, 13*, 257-283.

APPENDICES

Appendix 1: Student Questionnaire

IRB No. 2208/003-006 유효기간: 2023년 08월 17일

1학년 ( )반 ( )번

다음은 컴퓨터(디벗)와 코퍼스를 활용한 영어 유의어 수업에 대한 여러분의 솔직한 의견을 묻는 설문지입니다. 성적에 반영되지 않고 비밀이 보장되니 솔직하게 답해주세요.

A. 영어를 몇 살(혹은 몇 학년) 때부터 배우기 시작했나요? ( )

B. 영어권 국가 (미국, 영국, 캐나다 등)에서 살아본 경험이 있나요? □ 예 □ 아니오

'예'라고 답했다면 어느 나라에서 몇 개월 거주했나요? ( )에서 ( )개월

C. 컴퓨터(디벗)와 코퍼스를 활용한 수업에서 가장 도움이 된 점은 어떤 것이었나요?

D. 컴퓨터(디벗)와 코퍼스를 활용한 수업에서 가장 어려웠던 점은 어떤 것이었나요?

E. 다음을 읽고 해당하는 곳에 체크하세요.

		매우 그렇지 않다	그렇지 않다	보통 이다	그렇다	매우 그렇다
1	코퍼스를 사용하는 것은 유사한 단어의 의미 차이를 배우는 데 도움이 된다.	□	□	□	□	□
2	코퍼스는 단어가 문장에서 쓰이는 용법을 배우는데 도움이 된다.	□	□	□	□	□
3	코퍼스를 통해 영어의 실제 문장 예시를 직접 접하게 되어 좋은 기회였다.	□	□	□	□	□
4	코퍼스를 사용해 패턴을 찾는 것은 시간이 많이 걸리고 어려웠다.	□	□	□	□	□
5	어려운 영어 단어가 많아서 코퍼스를 사용하기 힘들었다.	□	□	□	□	□
6	코퍼스 데이터에 문장이 너무 많아서 학습하기 곤란했다.	□	□	□	□	□
7	대체로 코퍼스는 유용한 학습 자료인 것 같다.	□	□	□	□	□
8	코퍼스는 활용한 다른 유형의 활동도 해 보고 싶다는 생각이 들었다.	□	□	□	□	□
9	코퍼스를 사용해 스스로 영어를 학습할 수 있다.	□	□	□	□	□
10	앞으로도 영어 수업시간에 코퍼스를 이용하여 수업하는 것을 추천한다.	□	□	□	□	□
11	컴퓨터를 사용한 영어 수업은 흥미로웠다.	□	□	□	□	□
12	컴퓨터를 사용한 영어 수업은 단어를 학습하는 데 도움이 되었다.	□	□	□	□	□
13	컴퓨터를 사용하는 것에 익숙하지 않아 수업 내용을 따라가기 힘들었다.	□	□	□	□	□
14	컴퓨터를 사용함으로써 활동지를 완성하는 시간이 단축되었다.	□	□	□	□	□
15	컴퓨터는 영어 수업시간에 계속해서 활용되어야 한다.	□	□	□	□	□

F. 그 외 수업에 관한 생각이나 의견이 있다면 자유롭게 적어주세요.

수고하셨습니다.

Ver 1.1(2022.08.18.)

Appendix 2: Pretest Material

IRB No. 2208/003-006 유효기간: 2023년 08월 17일

1학년 ( ) 반 ( ) 번

이 시험은 더 나은 수업을 위한 사전평가로 성적에는 반영되지 않습니다. 각 문제를 읽고 정답에 체크해 주세요. 정답을 모르겠다면 '잘 모르겠다'에 솔직하게 표시해 주세요.

※ 각 문장의 괄호 안에서 알맞은 단어를 선택해 문장을 완성하세요.

1. He began to **(tell / speak)** very quickly. ☐ 잘 모르겠다
2. She **(says / talks)** she needs more time. ☐ 잘 모르겠다
3. Perhaps they **(heard / listened)** me shout. ☐ 잘 모르겠다
4. First, she had to **(end / finish)** her studies. ☐ 잘 모르겠다
5. The police are going to **(look / watch)** you for some time. ☐ 잘 모르겠다
6. Nobody **(looked / saw)** her when she fell down the stairs. ☐ 잘 모르겠다

※ 다음 빈칸에 공통으로 들어갈 단어를 고르세요.

7.

I could ( ) my voice.
He said he didn't ( ) anything.

① hear ② listen ③ say ④ tell ☐ 잘 모르겠다

8.

I'm going to ( ) you something.
I must ( ) this secret to someone.

① hear ② listen ③ say ④ tell ☐ 잘 모르겠다

9.

I must ( ) writing my new book.
I want to ( ) it, but it's very difficult.

① end ② finish ③ watch ④ look ☐ 잘 모르겠다

10.

We didn't ( ) anybody.
I don't want to ( ) you again.

① look ② see ③ end ④ finish ☐ 잘 모르겠다

※ 다음 밑줄 친 동사의 사용이 자연스러우면 O, 어색하면 X표를 하세요.
어색하다고 생각되는 경우 올바르게 고쳐보세요.

11. Anna listens carefully to her daughter. ( )
고친 문장: _____ ☐ 잘 모르겠다

12. They turned to watch at the house. ( )
고친 문장: _____ ☐ 잘 모르겠다

13. I want to say to you. ( )
고친 문장: _____ ☐ 잘 모르겠다

Ver 1.1(2022.08.18.)

14. Look! There he is! ( )
고친 문장: _____ □ 잘 모르겠다

15. Then summer ended and September came. ( )
고친 문장: _____ □ 잘 모르겠다

16. You can talk that the weather is nice. ( )
고친 문장: _____ □ 잘 모르겠다

※ 우리말과 같은 뜻이 되도록 문장의 빈칸을 완성해 보세요.

17. 마침내, 전쟁은 1814년에 끝났다.
Finally, the war _____ed in 1814.

18. 제발 내 말을 듣고 나를 도와줘.
Please _____ to me and help me.

19. 너의 이야기를 나에게 말해줘.
_____ me your story.

20. 그녀는 영화 보는 것을 사랑한다.
She loves to _____ movies.

Appendix 3: Posttest Material

IRB No. 2208/003-006 유효기간: 2023년 08월 17일

1학년 ( ) 반 ( ) 번

이 시험은 여러분이 배운 내용을 바탕으로 한 사후평가로 성적에는 반영되지 않습니다. 각 문제를 읽고 정답에 체크해 주세요. 정답을 모르겠다면 '잘 모르겠다'에 솔직하게 표시해 주세요.

※ 각 문장의 괄호 안에서 알맞은 단어를 선택해 문장을 완성하세요.

1. What did your mother **(say / tell)**? ☐ 잘 모르겠다
2. Do you all **(speak / talk)** English? ☐ 잘 모르겠다
3. Now **(hear / listen)** closely to the speech. ☐ 잘 모르겠다
4. When did you **(look / see)** her yesterday? ☐ 잘 모르겠다
5. You can sit and **(look / watch)** the birds here. ☐ 잘 모르겠다
6. It took 14 years to **(end / finish)** this painting. ☐ 잘 모르겠다

※ 다음 빈칸에 공통으로 들어갈 단어를 고르세요.

7.

I could see and ( ) more clearly.
He didn't ( ) the noise behind him.

① hear ② listen ③ say ④ tell ☐ 잘 모르겠다

8.

Let's ( ) this wonderful food.
Please ( ) your very interesting story.

① look ② watch ③ end ④ finish ☐ 잘 모르겠다

9.

I'll ( ) them the truth.
Let me ( ) you what happened to me today.

① hear ② listen ③ say ④ tell ☐ 잘 모르겠다

10.

Did you ( ) at my papers?
Oh, ( ), there, across the street!

① see ② look ③ end ④ finish ☐ 잘 모르겠다

※ 다음 밑줄 친 동사의 사용이 자연스러우면 O, 어색하면 X표를 하세요.
어색하다고 생각되는 경우 올바르게 고쳐보세요.

11. I want to <u>say</u> to you. ( )
고친 문장: _____ ☐ 잘 모르겠다

12. Then summer <u>ended</u> and September came. ( )
고친 문장: _____ ☐ 잘 모르겠다

Ver 1.2(2022.09.13.)

13. They turned to <u>watch</u> at the house. ( )
고친 문장: _____ ☐ 잘 모르겠다

14. I didn't listen to music or <u>see</u> television. ( )
고친 문장: _____ ☐ 잘 모르겠다

15. You can <u>talk</u> that the weather is nice. ( )
고친 문장: _____ ☐ 잘 모르겠다

16. The other girls loved to <u>listen</u> to her. ( )
고친 문장: _____ ☐ 잘 모르겠다

※ 우리말과 같은 뜻이 되도록 문장의 빈칸을 한 단어로 완성해 보세요.

17. 그녀의 목소리를 들었니?
Did you _____ her voice?

18. 내가 너의 사진을 봐도 될까?
Can I _____ your photos?

19. 이것에 대해 아무에게도 말하지마라.
Don't _____ anybody about it.

20. 1956년은 흑인에 대한 끔찍한 폭력과 함께 끝났다.
The year 1956 _____ed with terrible violence against blacks.

Appendix 4: Worksheet Samples

Middle School English 1	Finding the Words 'speak' and 'talk'	Student No.

☺ Let's find out the differences between the words 'speak' and 'talk' 동사 'speak'와 'talk'의 차이점을 찾아봅시다.
Remember! Everything is in lextutor.ca Use your own smart device. 각자의 디벗을 사용해 lextutor 사이트에 들어가 보세요.

Step 1. *Speak* means '( )' and *Talk* means '( )' 각 단어의 뜻을 적어보세요.

Step 2. Enter the word 'speak'. How many examples can you find? Equals = ( ) / Family = ( ) 총 몇 개의 예시가 있나요?

Step 3. Read the sentences and find the pattern. 검색한 문장들을 읽고 패턴을 찾아봅시다.

1) What words can you find before the word 'speak'? 동사 speak 앞에는 주로 어떤 단어들이 오나요? ( ) =주어: ( )

2) What words or expressions can you find after the word 'speak'? 동사 speak 뒤에는 주로 어떤 표현들이 오나요?

	(1)	(2)	(3)
Example 1			
Example 2			
Example 3			

➜ Plus! 동사 speak는 주로 (일방적으로 말하는 상황 / 서로 말하는 상황)에서 사용되기도 합니다.

Step 4. Enter the word 'talk'. How many examples can you find? Equals = ( ) / Family = ( ) 총 몇 개의 예시가 있나요?

Step 5. Read the sentences and find the pattern. 검색한 문장들을 읽고 패턴을 찾아봅시다.

1) What words can you find before the word 'talk'? 동사 talk 앞에는 주로 어떤 단어들이 오나요? ( ) =주어: ( )

2) What words or expressions can you find after the word 'talk'? 동사 talk 뒤에는 주로 어떤 표현들이 오나요?

	(1)	(2)	(3)
Example 1			
Example 2			
Example 3			

➜ Plus! 동사 talk는 주로 (일방적으로 말하는 상황 / 서로 말하는 상황)에서 사용되기도 합니다.

Step 6. What are the similarities and differences between the words ***speak*** and ***talk***? 동사 ***speak***과 ***talk***은 어떤 공통점과 차이점이 있나요?

Similarities 공통점	
Differences 차이점	Speak는 주로 Talk는 주로

Step 7. Fill in the blanks based on the pattern that you found. 오늘 찾은 내용을 바탕으로 빈 칸에 알맞은 단어를 써 문장을 완성해 보세요.

(1) Tylor can ( ) Korean very well.

(2) Do you want to ( ) about this problem with me?

(3) Sarah ( )s to her mother about her school life every day.

(4) You should ( ) loudly in front of other people.

Step 8. Make your own sentences using the words ***speak*** and ***talk***. 동사 ***speak***과 ***talk***을 각각 활용하여 자신만의 문장을 한 개씩 만들어 보세요.

수업 중 완성하는 활동지는 성적에 반영되지 않습니다. 끝까지 최선을 다해서 활동지를 완성해 주세요 ☺

Middle School English 1	Finding the Words 'listen' and 'hear'	Student

☺ Let's find out the differences between the words 'listen' and 'hear' 동사 'listen'과 'hear'의 차이점을 찾아봅시다.
Remember! Everything is in lextutor.ca Use your own smart device. 각자의 디벗을 사용해 lextutor 사이트에 들어가 보세요.

Step 1. *Listen* means '( )' and *Hear* means '( )' 각 단어의 뜻을 적어보세요.

Step 2. Enter the word 'listen'. How many examples can you find? Equals = ( ) / Family = ( ) 총 몇 개의 예시가 있나요?

Step 3. Read the sentences and find the pattern. 검색한 문장들을 읽고 패턴을 찾아봅시다.

1) What words or expressions can you find after the word 'listen'? 동사 listen 뒤에는 주로 어떤 표현들이 오나요?

	(1)	(2)
Example 1		
Example 2		
Example 3		
Example 4		
Example 5		

2) Translate the sentences you found. In which situation the verb 'listen' has usually been used?
찾은 문장을 해석해 보세요. 동사 listen은 주로 어떤 상황에서 사용된다고 생각하나요?

Example Sentences	Translation (해석)
동사 listen은 주로 (자연스럽게 들리는 상황 / 주의 깊게 들어야하는 상황)에서 많이 사용된다.	

Step 4. Enter the word 'hear'. How many examples can you find? Equals = ( ) / Family = ( ) 총 몇 개의 예시가 있나요?

Step 3. Read the sentences and find the pattern. 검색한 문장들을 읽고 패턴을 찾아봅시다.

1) What words or expressions can you find after the word 'hear'? 동사 hear 뒤에는 주로 어떤 표현들이 오나요?

	(1)	(2)
Example 1		
Example 2		
Example 3		

2) Translate the sentences you found. In which situation the verb 'hear' has usually been used?
찾은 문장을 해석해 보세요. 동사 hear은 주로 어떤 상황에서 사용된다고 생각하나요?

Example Sentences	Translation (해석)
동사 hear은 주로 (자연스럽게 들리는 상황 / 주의 깊게 들어야하는 상황)에서 많이 사용된다.	

Step 6. What are the similarities and differences between the words ***listen*** and ***hear***? 동사 listen과 hear은 어떤 공통점과 차이점이 있나요?

Similarities 공통점	
Differences 차이점	Listen은 Hear은

Step 7. Fill in the blanks based on the pattern that you found. 오늘 찾은 내용을 바탕으로 빈 칸에 알맞은 단어를 써 문장을 완성해 보세요.

(1) He said he didn't ( ) anything.
(2) She always ( ) carefully to her daughter.
(3) I didn't ( ) the noise behind him.
(4) We love to ( ) to his story.

Step 8. Make your own sentences using the words ***hear*** and ***listen***. 동사 *hear*과 *listen*을 각각 활용하여 자신만의 문장을 한 개씩 만들어 보세요.

Today's Learning Log ☺

수업 중 완성하는 활동지는 성적에 반영되지 않습니다. 끝까지 최선을 다해서 활동지를 완성해 주세요 ☺

Appendix 5: Pre-selected Interview Questions

Interview Questions	
Session 1	Have you heard about corpus?
	What was the most impressive thing about (using) corpus?
	How was today's activity? Was it easy for you to find the similarities and differences between the given words?
	Were there any difficulties of using your own computer?
Session 2	How was today's activity? Was it easy for you to find the similarities and differences between the given words?
	Could you finish today's worksheet on time? If not, what was the reason?
	Compared to the last session, what becomes easier or harder to do?
Session 3	How was today's activity? Was it easy for you to find the similarities and differences between the given words?
	We compared three words at once for the last activity. Wasn't it difficult to finish on time?
	Today was the last session of discovering patterns of synonymous verbs. In what aspect do you think you made a progress?
Session 4	What was your impression about corpus-based English vocabulary learning?
	Do you have any other things you want to discover and learn using corpus?
	Can you think of any advantage or disadvantage of DDL?
	Can you think of any advantage or disadvantage of CALL?

Appendix 6: Pretest and Posttest Scores of the Participants

ID	Pretest Score	Posttest Score
1	7	16
2	11	15
3	18	20
4	19	18
5	14	18
6	14	14
7	0	4
8	5	14
9	17	19
10	3	5
11	10	19
12	7	14
13	7	13
14	12	11
15	7	10
16	7	9
17	17	18
18	15	18
19	11	16
20	13	12
21	4	14
22	14	17
23	16	14
24	16	16
25	3	4
26	8	12
27	12	17
28	15	19
29	4	10

국 문 초 록

영어를 학습하는 데에 있어 단어를 이해하는 것은 의사소통 역량의 핵심 요소 중 하나이다. 하지만 일반적인 단어들과는 달리 영어를 외국어로 학습하는 학습자들은 모국어의 뜻은 같으나 미묘한 차이를 내포하는 유의어 학습을 특히 더 어려워하는 경향을 보여왔다. 본 연구에서는 학습자가 직접 온라인 코퍼스 자료에 접근하여 언어 사용의 특징과 규칙을 발견하는 언어자료기반학습(Data-Driven Learning; DDL)을 유의어 학습의 방법으로 제시한다. 선행 연구를 살펴보았을 때, 언어자료기반학습은 학습자의 영어 능숙도에 따라 그 효과가 달라질 수 있기 때문에 대체로 성인 영어 학습자를 위주로 연구되어 왔고, 다양한 기술적·환경적 제약으로 인해 편집된 코퍼스 자료를 기반으로 한 것을 확인할 수 있었다.

따라서 본 연구는 컴퓨터를 활용한 언어자료기반학습이 한국인 중학생의 영어 유의어 학습에 미치는 영향과 효과성을 검증한다. 구체적으로, 서울특별시교육청이 2022학년도부터 실시하는 스마트기기 휴대 학습 「디벗」을 활용하여 중학교 영어 학습자들이 코퍼스 자료로부터 스스로 활용 예시를 찾아 영어 유의어 간의 차이를 발견할 수 있는지 살펴보고 이 과정을 통해 학습자의 태도, 동기 및 학습에서의 자율성의 변화를 알아보고자 한다.

연구를 위해 서울의 한 중학교에 재학 중인 1학년 학생 29명을 모집하였다. 학생들은 사전 평가와 실험 진행을 위한 한 차시의 교육을 완료한 후 3주 간 총 다섯 쌍의 영어 유의어 간의 차이를 스스로 발견하는 시간을 가졌다. 모든 수업이 끝난 후 사후 평가를 실시하여 사전, 사후 결과를 비교했으며 추가적으로 학생들이 실시한 설문지와 사전 동의를 받은 3명의 학생과 개별 면담 결과를 분석하였다.

분석 결과 다음과 같은 컴퓨터 활용 언어자료기반학습의 긍정적인 효과와 학습자에 미치는 영향이 나타났다. 첫째, 인지적인 측면에서 컴퓨터 활용 언어자료기반학습은 유의어 차이의 학습에 효과적인 것으로 드러났다. 둘째, 정의적인 측면에서 중학교 영어 학습자들은 컴퓨터 활용 언어자료기반학습에 긍정적인 인식을 보이는 것으로 드러났다. 특히, 학습자들은 직접 컴퓨터를 활용하여 수업에 참여한다는 점에 흥미를 느끼는 경향을 보였다. 하지만, 개별 면담 결과를 통해 학습자간 차이가 있다는 점이 발견되었다. 상위권 학습자는 컴퓨터 활용 언어자료기반학습을 통한 유의어 학습에 어려움 없이 참여하였지만, 중위권 학습자는 때때로 교수자의 비계를 필요로 하였으며, 하위권 학습자의 경우 학습 과정 및 방법에 적응하는 데 상대적으로 긴 시간이 걸리고 과제 완성에 어려움을 느껴 교수자의 많은 설명과 비계가 뒷받침되어야 함을 시사했다.

적은 표본 크기 및 방법론적 한계에도 불구하고, 본 연구는

컴퓨터 활용 언어자료기반학습이 유의어 학습의 효과적인 방법임을 증명함과 동시에 한국 중학생 학습자들의 영어 수업 참여도와 자기주도적 학습 역량을 성공적으로 증진시켰다. 더 나아가 본 연구의 결과는 언어자료기반학습이 한국의 중학교 영어 수업에 어떻게 적용될 수 있는지, 그리고 코퍼스가 중학생들의 영어 학습 도구로 어떻게 사용될 수 있는지에 대한 통찰을 제공하는 영어 교육학적 함의를 지닌다.

주요어: 코퍼스, 코퍼스 기반 언어학습, 자료기반학습(DDL), 컴퓨터 보조 언어학습(CALL), 영어 유의어, 어휘학습

학 번: 2021-20625