

EFL Teachers' Perceptions of Data-Driven Learning (DDL) in L2 Grammar Teaching

Supakorn Phoocharoensil

*Language Institute Thammasat University
supakorn.p@litu.tu.ac.th*

Abstract

Data-Driven Learning (DDL) has been shown to have pedagogical advantages in English language education due to its ability to promote learner autonomy through inductive learning based on corpus data. This study aimed to investigate in-service English teachers' and university lecturers' perceptions of learning English conditionals through DDL materials in a computer-based environment in which learners interact with naturally-occurring L2 data from the Corpus of Contemporary American English (COCA) for six weeks. Four main themes arising from the interview results showed mixed opinions regarding the benefits of DDL, with positive responses outweighing the negative. DDL helps expand learners' grammatical knowledge, allowing them to notice authentic mixed-types of conditionals existing in English and encourages students to learn L2 grammar rules through self-discovery. However, some participants viewed DDL as a challenge as both teachers and learners require ICT skills and specialized knowledge to work with corpus data effectively. Furthermore, school policies and budget concerns are obstacles to DDL implementation as learners need technical support from schools.

Introduction

In the acquisition of a second language (L2), it is vital that learners do not ignore the importance of grammar, as it is considered fundamental to success in L2 communication. L2 English grammar can be challenging for learners aiming for correct usage. Two main perspectives on grammar are prescriptive, which concentrates on strict, traditional rules regarding the correctness of language use, and descriptive, which deals with probabilistic rules reflecting authenticity in language (Timmis, 2015). The current trends in language description in relation to grammar instruction have been influenced by corpus-based research, which reveals a number of interesting facts about the way native speakers of English use their language in real life. However, grammar lessons in traditional textbooks often ignore these authentic occurrences (Jones & Waller, 2015).

Of the corpus-aided language learning approaches that have been on the rise over the past decade, one notable approach known as

Data-Driven Learning (DDL) in English Language Teaching (ELT) has been the subject of a number of empirical studies devoted to examining its effectiveness (e.g., Geluso & Yamaguchi, 2014; Karras, 2016; Larsen-Walker, 2017; Rezaee et al., 2015; Yamtui & Phoocharoensil, 2019). DDL prompts learners to take on the role of language researchers through noticing and analyzing designated linguistic features, which will then lead to generalization of L2 rules (Lin & Lee, 2019). The entire process of DDL involves three main steps, namely, *identify*, *classify*, and *generalize* (Johns, 1991). This process closely reflects the nature of discovery and enhances DDL learners' awareness of L2 linguistic items. DDL is unique in that it not only supports the noticing hypothesis in second language acquisition and empowers learners to develop linguistic competence in L2 (Flowerdew, 2012), but its use of corpus-based material and activities serve to interest language learners due to the richness and authenticity of the L2 data (Lin & Lee, 2019). The pedagogical effect of DDL is confirmed by positive empirical

investigations into DDL students' learning outcomes and attitudes.

Among the DDL studies available thus far are those experimenting on the use of DDL in English grammar teaching (e.g., Boontam & Phoocharoensil, 2018; Koosha & Jafarpour, 2006; Kumpawan & Nishigaki, 2020; Petcharinphan, 2020; Rapti, 2010). Although there have been studies on the use of DDL in grammar instruction in Thai EFL contexts at the primary (e.g., Boontam & Phoocharoensil, 2018), secondary (e.g., Petcharinphan, 2020), and tertiary levels (e.g., Dankittikul & Laohawiriyanon, 2018; Sripicharn, 2002), it is necessary to consider introducing DDL to in-service English teachers in order to raise their awareness of the significance of DDL in ELT and encourage them to incorporate DDL activities in their classrooms. The present study therefore aimed to investigate Ph.D. (ELT) students' perceptions of DDL after providing them with the basic principles and applications of DDL with an emphasis on grammar teaching through an intensive 6-week workshop. The participants were asked to reflect on their experience of corpus consultation with regard to English conditionals and estimate the potential benefits of applying DDL activities in their future instruction.

Literature Review

Language Corpora and ELT

A *corpus* (*corpora*, plural) has been defined as a collection of digital written or spoken texts stored on a computer (O'Keeffe et al., 2007). A more specific definition, found in Corpus Linguistics (CL), states that a corpus refers to a collection of electronic naturally-occurring texts systematically collected. Prior to the widespread use of computers, the electronic text collection had to be manually acquired. The advent of more advanced computers and the digitization of spoken and written language has allowed for a more systematic corpus-based collection of language data (Friginal, 2018). Thus, a language corpus is now computerized and searchable by computer software. Not only do language researchers apply corpus techniques to empirical

investigations but also EFL/ESL teachers find corpora beneficial for authentic-materials to use in their lessons (Jones & Waller, 2015; Szudarski, 2018).

Information from language corpora informs and promotes ELT in a number of ways. For instance, learners can investigate actual language use and variation using the empirical data from L1-English speaker corpora, e.g., the Corpus of Contemporary American English (COCA) (Crawford & Csomay, 2016), while teachers can describe and introduce L2-English high-frequency grammatical patterns (Timmis, 2015). In the context of ELT, CL methods provide learners with relevant, meaningful data, e.g., authentic language examples, frequency and distributions of English grammatical and lexical patterns, etc. As Friginal (2018) notes, “[m]obile technology, individualized instruction, and big data visualization as integral parts of CL all contribute to how digital learners may, in fact, fully adapt and appreciate corpus-based approaches in their learning of English.” (p.6). Based on corpus-informed results derived from a large amount of data, ELT researchers and practitioners can enhance generalizability and validity of the linguistic phenomena of their own interest and then incorporate the findings into their classrooms.

Data-Driven Learning (DDL) and Grammar Teaching

From the mid-1980s to the early 1990s, corpus-based language teaching was viewed as innovative though largely inaccessible to English teachers (Friginal, 2018). One of the most well-known CL-based inductive teaching approaches is Data-Driven Learning (DDL), which was introduced to the language classroom by Tim Johns (1991). The DDL approach is known as ‘Classroom Concordancing’, which encouraged learners to study language samples extracted from concordances, i.e., numerous corpus instances of a particular word or phrase presented in context, and make observations on its meaning, use, and grammatical patterns based on the language evidence. As Johns proposed, DDL learners are expected to work on linguistic data

as if they were language researchers or detectives in the typical learning process of ‘identification-classification-generalization’, i.e., a pedagogical formula prescribed by Johns (1991). The main purpose of DDL is to familiarize learners with corpus tools and resources and then train them to use corpora in their L2 learning. DDL is a powerful teaching approach that “expose[s] learners to authentic language in order for them to examine and understand how the language is structured and used naturally – leading, consequently, to their own successful use” (Friginal, 2018, p. 42). As a result, learner autonomy is boosted through learners’ self-discovery of L2 rules.

Regarding the function of DDL in the process of learners’ target language rule discovery, it is important for language teachers to equip learners with a sufficient context through which learners can develop their L2 knowledge (Timmis, 2015). Having been exposed to corpus data, learners are encouraged to formulate hypotheses from the data with guidance from teachers who will ask questions to check learners’ understanding and acquired knowledge. In observing L2 usage and pattern frequency in the responses, learners have to become well-trained in the use of the related corpus software, e.g., AntConc, Wordsmith, or skilled in searching already-existing corpora, e.g., COCA, BNC. Teachers will then give feedback on whether the questions have been correctly answered, and learners will expand their new L2 knowledge which will then be integrated into what is already known (Boulton, 2009). To sum up, DDL is the guided exploration of L2 features and patterns. The discovery of L2 patterns allows learners to “...articulate insights and experience a degree of self-sufficiency in their language learning.” (Friginal, 2018, p. 40).

The popularity of DDL as a pedagogical tool has increased in ELT contexts as it primarily incorporates register awareness and grammar/vocabulary instruction directly based on language data. According to Hadley (2002), DDL learners can access a great amount of authentic data using corpus software programs, thus allowing them to observe common naturally-occurring phraseological patterns in language

samples. DDL is basically a form of consciousness-raising of L2 grammar. DDL teachers should keep in mind that the rationale for DDL deals with authenticity and autonomy. In terms of authenticity, the language data in the materials provided for L2 learners should be representative of real language use as asking learners to look at corpus data can result in language feature descriptions that are more accurate and practical than those found in ELT materials in general (Friginal, 2018; O’Keefe et al., 2007). For example, while most traditional ELT textbooks mainly present only four major types of English conditionals, namely the three classic types and Zero conditional, many corpus-based studies have demonstrated that there actually exist several types of conditionals known as ‘mixed’ or ‘alternative’ conditionals in native-speaker English corpora, e.g., BNC (Jones & Waller, 2011), written BNC or BNCw (Gabrielatos, 2019, 2021), ICE-GB or International Corpus of English-Great Britain (Lastres-López, 2020), COCA (Phoocharoensil, 2014), and POTTI or Post-Observation-Teacher-Training Interactions (Farr & McCarthy, 2002).

Some examples of corpus-based alternative conditionals in BNCw and COCA are given below:

(1) It probably *didn't cover* the petrol, if she *sat* down to work it out, which she *didn't*. (BNCw)

(Gabrielatos, 2021, p. 104)

(2) If he *won't tell* the truth, I *was going to tell* the truth. (COCA)

(Phoocharoensil, 2014, p. 69)

A number of studies have revealed considerable benefits of DDL in the improvement of students’ L2 English grammar, as well as learners’ positive attitudes towards DDL, e.g., general grammar (e.g., Rapti, 2010), auxiliaries (e.g., Kumpawan & Nishigaki, 2020), prepositions (e.g., Boontam & Phoocharoensil, 2018; Koosha & Jafarpour, 2006; Kumpawan &

Nishigaki, 2020), gerunds and infinitives (e.g., Kumpawan & Nishigaki, 2020), and logical connectors (Petcharinphan, 2020).

In an experimental study, Rapti (2010) examined the application of DDL in teaching common English grammar items and patterns to adolescent L1-Greek students. Concordance-based grammar materials were introduced to the experimental group, while the control group was taught using a conventional grammar book. The findings showed that the participants in the experimental group improved noticing skills through DDL lessons, despite some difficulties regarding unknown vocabulary as well as the concordance format, which underscores the need for teacher facilitation. As shown in the questionnaire and interview results, the participants recognized the contribution and potential of corpora, but there were variations in motivation levels when studying grammar through DDL. When comparing the test performances of the two groups, more participants of the experimental group had higher scores than those of the control group. Therefore, both the DDL effectiveness in grammar instruction and the students' positive attitudes towards this method can be confirmed.

Koosha and Jafarpour (2006) conducted an experimental study on the effectiveness of DDL on Iranian EFL learners' development of collocations of prepositions. They collected data from 200 Iranian English-major students equally assigned to three control groups and three experimental groups differing in L2 English proficiency levels. The students in the control group were taught using a conventional approach in which prepositions and their collocational patterns were explicitly introduced to them in English or Farsi, whereas those in the experimental groups received DDL instruction based on concordance lines from the Brown Corpus Online illustrating the six target collocational patterns, i.e., adjective + preposition, preposition + noun, noun + preposition, verb + preposition, preposition + preposition collocations, and idiomatic expressions. As shown in the results, the students in the DDL groups outperformed those in the conventional ones, showing that DDL instruction

gave students an advantage. This furthermore suggests that DDL enables learners to learn L2 collocations of prepositions with a greater success.

Boontam and Phoocharoensil (2018) also confirmed the effectiveness of DDL in building learners' knowledge of L2 English prepositions. The study used a paper-based data-driven learning (DDL) method in developing Thai A2 grade-4 students' knowledge of three English prepositions, namely, *during*, *among*, and *between*. The mean scores of the post-test, comprised of a gap-filling test, a grammaticality judgement test, and a sentence-building test, indicated a statistically significant improvement in the learners' preposition knowledge after they had completed 6-weeks of DDL lessons. The questionnaire and interview results also point out the students' positive attitudes towards preposition learning through the DDL activities, which many found interesting and fun. A majority indicated they wanted to study other English lessons through DDL in the future as they considered DDL as a useful resource for acquiring new grammatical knowledge.

There has also been an effort in viewing the way DDL promotes L2 grammar and writing, i.e., a skill relative to grammar. Kumpawan and Nishigaki (2020) conducted an experimental study and found a relationship between DDL and A1 grade-10 Thai EFL learners' L2 grammatical knowledge and writing skill acquisition. Focusing on gerunds/infinitives, the present form of the auxiliary *have*, and the prepositions of time *in*, *on*, and *at*, this study reported on statistically significant differences between a pre-test and a post-test and between a pre-test and a delayed post-test, which implied the effectiveness of DDL in the learners' production and retention of the target grammatical forms. In other words, they were able to notice the target grammar and then improve their writing skills. The questionnaire data showed the students' satisfaction with DDL at a moderate level, suggesting that they were not familiar with DDL-inductive learning and needed more time and training to shift from a passive to more active learning style.

DDL also proved effective in the introduction of logical connectors. In Petcharinphan's (2020) study, the effects of the integrated inductive approach between grammar-consciousness raising (GCR) task and DDL in increasing mixed-and-low proficiency grade-12 Thai EFL learners' logical connector knowledge were investigated. For the experimental group of 30 students, the integrated inductive approach on the GCR task and DDL was implemented in five lesson plans with ten different logical connectors, i.e., *due to*, *in order to*, *as a result*, *as a result of*, *in contrast to*, *on the contrary*, *despite*, *whereas*, *in addition to*, and *as well*, whereas the same logical connectors were taught through deductive lessons in the control group. The findings demonstrated that the experimental group, implemented with the integrated inductive approach scored higher than the control group implemented with the deductive approach. In other words, DDL appeared to assist the students who received the DDL treatment to learn logical connectors. In addition, those in the experimental group had positive attitudes toward the integrated inductive approach.

Despite the fact that DDL has proven useful in L2 learning as exhibited in many studies, some reservations about DDL should be taken into consideration. From learner perspectives, learning through DDL can be complicated and even intimidating particularly for low-level language learners. Training learners to become grammatically conscious through noticing patterns of L2 can be a challenge as it is time-consuming and may cause learners to become confused and frustrated when flooded with a vast amount of unorganized data (Friginal, 2018). As Timmis (2015) remarks, changing learners from passive recipients of L2 knowledge to active, capable observers requires adequate practice. It is implied that teachers' advice and close supervision is of paramount importance in DDL classrooms. Furthermore, some DDL teachers expressed concern over their implementation of DDL activities. Teachers need to adapt teaching approaches and acquire digital literacy so that they will master investigation of corpus data and be able to produce DDL-based materials for their instruction (Römer, 2011). It is worth noting that it usually takes teachers weeks to properly

introduce and use corpus tools in teaching (Friginal, 2018). Geluso and Yamaguchi (2014) maintain that DDL requires both learners' and teachers' investment both in time and effort. Weeks should be spared to give students a clear demonstration of how to effectively search for and analyze data from concordance lines. This is to ensure that students will be convinced of the authenticity of corpus data and motivated enough to conduct their own exploration of DDL activities.

This study was carried out to respond to the research question below:

Research question: What are the perceptions of Thai EFL in-service teachers at secondary and tertiary levels towards the use of DDL in teaching conditionals?

Methodology

Participant Selection

This study used convenience sampling in recruiting the twelve Thai participants, all of whom were enrolled in a doctoral elective course in Corpus Linguistics and ELT at a public university in Thailand. While eight of them were teaching at the university level, the other four were secondary English teachers. All the participants had been learning English as a foreign language for at least 18 years, and their teaching experience ranged from 3 to 8 years when the data collection took place. Based on the scores of the standardized English test, i.e., an in-house one equated with TOEFL-iBT, the Ph.D. students were all considered upper-intermediate and advanced English learners whose scores corresponded to the B2 to C1 levels of the Common European Framework of Reference for Languages (CEFR). The high-proficiency in-service English teachers were targeted for this study due to two main reasons. First, DDL learners with high English competency tend to be more successful in noticing L2 patterns shown in concordance lines and drawing more acceptable generalizations of the related L2 rules in comparison with those of low proficiency, who

are often confronted with difficulties in corpus data interpretation (Flowerdew, 2012). Second, as in-service teachers with experience in teaching English grammar, the participants had more opportunities to relate or apply the DDL concepts and activities to their own instruction in the future. Thus, it was appropriate to have them assess the advantages of DDL. Having been informed in advance that they could choose not to join the study or even leave at any time of the experiment and interview, the participants all agreed to take part in the study, showing their enthusiasm in completing all the DDL tasks during the experiment and responding to all the interview questions.

Data Collection Procedure & Data Analysis

This study used a one-group quasi-experimental design as it did not involve randomized sample selection. The study was part of a 16-week Ph.D. course in Corpus Linguistics and ELT, and 7 weeks (21 hours) were dedicated to a DDL workshop, in which the first week (3 hours) was spent on special training in exploring data from the Corpus of Contemporary American English (COCA) for DDL. The experiment was conducted over a six-week period, equal to 18 hours, addressing the English conditional. The conditional was selected not just because it concerns complicated grammar rules that frequently cause confusion for learners (Cowan, 2008) but also because there are some alternative usages that differ from what is usually found in ELT textbooks, and such authentic usages or patterns can be observed through corpus data. The three DDL steps were introduced to the participants, namely, 'identification', 'classification', and 'generalization' (Johns, 1991). More precisely, they learned to notice and identify the frequent grammatical patterns of conditionals and relative clauses from the searched concordance lines before they continued to categorize the observed patterns based on the corpus data. Finally, the participants practiced hypothesis formulating and testing skills to inductively construct L2 syntactic rules and patterns presented through the concordance lines.

Of the six-week experimental period, the first three weeks were devoted to DDL lessons on the three classic types of conditionals, whereas the next three focused on the zero conditional and alternative or mixed types. The researcher decided to use the classic computer-based DDL on the grounds that the participants were adult high-proficiency English teachers and could effectively conduct their own searches in COCA in the DDL process. The workshop was conducted via Zoom videoconferencing due to the COVID-19 pandemic. In Weeks 1-3, the use of COCA data to learn the fundamental concepts of the target grammar point was demonstrated to the participants. It is worth noting that the participants, in contrast to beginning learners, were teachers who possessed a very good understanding of conditionals, as revealed by the preliminary, informal interview results. Because of this, the workshop trainer, i.e., the researcher, was able to introduce the means of consulting COCA in response to grammar tasks given. The participants were also taught how to prepare paper-based DDL materials in case their future DDL students' access to real-time corpus data was limited or unavailable. The DDL tasks in Weeks 1-3 were targeted at familiarizing students with discovering the basic rules governing the usages of both grammar points. The DDL tasks for Weeks 5-6 were designed to raise the students', as well as the in-service teacher participants', awareness concerning the unusual patterns or occurrences, i.e., mixed types of conditionals, which is infrequently found in most ELT textbooks.

After the six-week period of DDL introduction, the participants were individually asked to join a semi-structured online interview via Zoom videoconferencing to determine their opinions on the experience, benefits, and challenges of learning grammar through DDL. The interviews were conducted in Thai as this was the native language of the researcher and the participants and eliminated any language problems that might occur during the data collection process. They were audio-recorded with the participants' permission and then transcribed. In addition to consent forms which the participants completed, their confidentiality, right to withdraw, and anonymity, were ensured.

Thematic content analysis was used to analyze the transcribed data. To achieve intercoder reliability, a qualified intercoder holding a Ph.D. in ELT was invited to code the data based upon the developed coding scheme. Both coders' percentage of agreement was approximately 85%, confirming the consistency and intercoder reliability.

Results and Discussion

Based on the interview results, the participants' responses to the DDL approach to grammar teaching were classified into four main themes: 1) Grammatical Knowledge Expansion with DDL, 2) Learner Autonomy Promotion, 3) Digital Literacy and Technical Challenges, and 4) Budget and Education Policy.

Theme 1: Grammatical Knowledge Expansion with DDL

Overall, most of the participants, including eight university lecturers and two secondary school teachers reported favorable opinions towards the DDL training. Those with positive attitudes towards teaching grammar through DDL were of the view that they could incorporate DDL tasks and activities into their own grammar lessons, which will help promote students' inductive learning through noticing skills and self-discovery of L2 rules. Three participants remarked that DDL helped broaden their understanding of mixed types of English conditionals as they had previously been aware of only three conventional types of if-conditionals and the Zero conditional, i.e., the type used to refer to scientific facts or something that is generally true (Larsen-Freeman & Celce-Murcia, 2015). To their surprise, as exemplified below, they admitted to learning the existence of alternative conditional types in American English through COCA data, i.e., the types that they would mark incorrect as they appeared different from the traditional patterns with which they were familiar, e.g., *If + present perfect + present simple*, *If + S + will + V.infinitive + S + was going to + V.infinitive*, or *If + past simple + present simple*. In actuality, such alternative

types are found in authentic English as revealed by a number of corpus-based studies (e.g., Gabrielatos, 2019, 2021; Jones & Waller, 2011; Phoocharoensil, 2014).

Participant A (tertiary)

I never knew there are other possible types of if-conditionals apart from what I see in general grammar references.

Participant D (tertiary)

This is an eye-opening experience. Using corpora to learn grammar allows me to access mixed types of conditionals. I was not aware of them existing in English and I think I should introduce this interesting fact to my students when teaching conditionals.

Participant I (secondary)

If I hadn't been introduced to these mixed types, I would think they were all wrong. Of course, I will definitely correct my students' use of such unconventional combinations of if-clause and main-clause patterns.

Furthermore, two participants admitted that in their teaching of Type-2 conditionals, which are used when referring to an unrealistic or hypothetical situation, they emphasized that students had to strictly conform to the prescriptive rule stating that *were* is the only single correct form of *be* regardless of the number of the subject(s). It came as a surprise to the in-service teachers to be informed by corpus data that not only is the form *was* allowed in English but *was* also occurs with far higher frequency than *were*, implying that the traditional grammar rules which they follow need reconsideration or revision.

Participant B (tertiary)

When I was learning English in High school, I was taught to always use 'were' in the second conditionals no matter what the subject is. And this is what I always keep in mind when I teach conditionals. Learning from corpus data has really widened my world, permitting me to know was is also possible.

Participant J (secondary)

It was a WOW experience to know corpus data really helps! The rule has changed. Now 'was' is acceptable and even more common than 'were'. I think I should teach this to my students so that they will know the contemporary usage of conditionals.

From their statements, it can be seen that the participants found DDL and corpus data useful and conducive to learning innovative or unorthodox patterns existing in authentic native speaker English. The fact that language corpora provide learners and teachers with new insights into English grammar found support in previous studies that confirmed the positive viewpoints of corpora and DDL. The results of the present study, in line with Boulton (2008), indicated several advantages of DDL tasks as corpus-based lessons can promote inductive learning and arouse students' interest to explore and make observations of language patterns that reflect practical and common usage. As noted by Boulton, learners found it interesting to be exposed to 'the language of today' from the contexts and concrete examples presented in concordance lines. In a similar vein, Boontam and Phoocharoensil (2018) also demonstrated that students who learned prepositions through DDL significantly improved their L2 knowledge of the target prepositions; moreover, 9 out of 28 students viewed DDL as a helpful resource to discover new patterns and structures. The discovery of unknown grammatical rules is in agreement with a study by O'Keefe, Carter & McCarthy (2007), in which they posited "[a] corpus can tell us different things about grammar. It can extend our understanding of traditional grammatical notions and categories..." (p. 100).

Theme 2: Learner Autonomy Promotion

All 12 participants considered DDL a very effective grammar teaching method that can promote learner autonomy to a great extent. As in-service teachers and university lecturers with experience in grammar instruction, they were convinced of the efficacy of DDL in strengthening students' ability to notice and classify patterns before they can finally generalize the related grammar rules based on the data from corpora. It is interesting to note that the participants would like to apply a DDL approach to their own grammar classrooms as they are of the opinion that DDL will increase students' confidence in learning and using L2 English skills, which corresponds to Crosthwaite et al. (2021) in that around 82% of primary and secondary English teachers in their study recognized the importance of DDL and agreed that corpora and DDL would be very useful resources for their teaching. Some participants' comments are as follows:

Participant C (tertiary)

DDL will make my students engaged in the classroom activities because it is fun and keeps them attentive at all times.

Participant H (tertiary)

I think my students majoring in English will find DDL and corpora fun and very useful. They will learn through their own trials and errors.

Participant F (tertiary)

Very exciting and thought-provoking activities! I am sure my students will enjoy learning new stuff in their own way.

Participant K (secondary)

I had a great time learning how to search for my new knowledge using corpora. My students,

although they are still young and their competency is not high, should become active, autonomous learners if introduced to DDL.

According to some participants, language corpora are useful references for grammar teaching as they provide a number of examples that facilitate students' L2 acquisition. They are of the opinion that the key to success in grammar learning lies in examples of real English, which allow for better understanding of usage of grammatical patterns in contexts. Corpus-based examples are clearer and more accurate than the examples they themselves create. Due to the fact that examples from native speaker corpora, e.g., COCA or BNC, are reliable and representative of authentic language use, learners will be certain of the new knowledge of L2 rules they learn from concordance lines. This is in line with Lin (2019) and Crosthwaite et al. (2021), who also reported on the benefits of DDL since students can interact with corpus-based material and become the 'host' of the class rather than mere passive learners as in traditional methods, e.g., Grammar Translation. DDL can promote the role of teachers as facilitators, which is consistent with Petcharinphan (2020), who showed that a student-centered grammar lesson encouraged students to explore and learn from corpus data; in addition, DDL permitted them to communicate in L2 English more than the previous traditional grammar lessons.

Theme 3: Digital Literacy and Technical Challenges

While many participants deemed DDL superior to traditional approaches to improving EFL students' grammar skills, they also expressed some reservations regarding utilization of DDL. Three in-service secondary teachers voiced their concerns about incorporating DDL materials into their grammar lessons to teach low-proficiency students, admitting that they were afraid their students could face difficulties in formulating rules from even concordance lines in paper-based DDL materials. Their concerns corroborated many past studies in that DDL-integrated materials, though useful in

constructing active, autonomous language learners, are often seen as difficult for low-proficiency learners (Flowerdew, 2012). In agreement with the secondary teachers, those teaching in the tertiary level also worried about using DDL in grammar instruction since university students also need time and guidance from their teachers in searching for the target grammar in corpora.

Participant I (secondary)

Secondary students may think working with corpora is difficult because they need some special training, as well as time to practice interpreting data from corpora.

Participant L (secondary)

Let alone students. A lot of teachers in Thailand are not tech-savvy and then refuse to use technology like corpora in teaching. This way they won't be able to give good advice to students in learning through DDL.

Participant E (tertiary)

Even though looking at and learning grammar from a DDL lesson could be fun, some students may not enjoy this teaching method. Some can be even confused or not satisfied with DDL, compared with a traditional approach, such as a deductive one, where rules are directly taught.

Participant G (tertiary)

Both students' and teachers' digital skills have to be ready for an implementation of DDL in grammar teaching.

The participants' reluctance to apply DDL to their instruction accords with Crosthwaite et al. (2021) in that there are technical challenges and barriers in implementing DDL in teaching practice. In the Indonesian EFL contexts, a

teaching method that involves computer-assisted language learning (CALL), including DDL, is largely problematic due to a lack of available devices or policies preventing the use of devices in the classroom, a lack of internet connectivity, a lack of space for CALL within the curriculum, and a lack of Indonesian teachers' sufficient technical knowledge. Likewise, the use of DDL in the ELT context of Taiwan was regarded as unfeasible as revealed in Lin's (2019) study. Negative responses and uncertainty from some of the primary and secondary teachers participating in that study included statements and questions such as "This method seems impossible for Taiwan", "it is too ideal a practice", and "Can Taiwanese students learn with it?". Working with DDL-based materials, students' mastery of corpus data analysis relies on strong support and close supervision from teachers (Friginal, 2018; Römer, 2011). The teacher as a facilitator is another key factor contributing to the success of DDL (Boontam & Phoocharoensil, 2018).

Theme 4: Budget and Education Policy

The findings showed that two schoolteachers mentioned some potential problems of DDL implementation in the context of Thai school education. Perceived difficulties in using computer-based version of DDL pertains to the availability of computer rooms where an entire class can search from corpora simultaneously. In addition, although it is now much easier for students to use hi-tech internet-connected devices, such as tablets and smartphones, it is impossible for every student to have access to personal electronic devices, which can be expensive for poor students in public schools. Access to the full version of many corpora, e.g., BNC, is not free of charge, or limited, e.g., COCA, which can impede their self-reliant online study.

Participant I (secondary)

I know learning L2 English grammar through DDL is great but how students from poor families can possess their own smart devices for online learning is still a big question.

Participant K (secondary)

The idea of DDL is good but hard in practice. Students at several Thai schools are not ready for corpus-based learning based on digital devices. Some schools don't even have enough computers for all students.

Another problem is attributed to the policy of some schools which prohibits students from using electronic devices like smartphones in class for fear that they could be distracted. This is in line with Crosthwaite et al. (2021), who reported on this problem in the Indonesian school context; that is, teachers wished their individual students had access to devices and high-speed Internet. Unfortunately, apart from the shortage of budget as mentioned earlier, finding a space for DDL, as well as other CALL activities within the Indonesian curriculum appears unfeasible since digitally-based education rarely gains support at the school and curricula levels in Indonesia. Paid access to the unlimited version of COCA has been noted as a limitation of in-class DDL by Yamtui and Phoocharoensil (2019), which is in consonance with the present-study teacher participants' concern.

Conclusion

This study has exhibited Thai EFL secondary teachers' and university lecturers' perceptions of DDL in grammar instruction. The target grammar focused on in this study was the conditional. From a 6-week training workshop, study participants were shown how to conduct searches in COCA and apply DDL-based activities to their grammar classrooms. The qualitative results based on a semi-structured interview showed the English language educators' positive viewpoints on teaching grammar through DDL-integrated tasks due to the fact that DDL can open up their worldviews on English grammar usage through authentic corpus data representing native speaker English. Not only did they claim to learn new uses of English conditionals, e.g., alternative types, but they also saw the considerable benefits

of DDL in promoting learner autonomy, which enhances learners' skills in noticing new and frequent patterns and formulating the rules from the corpus data. Nonetheless, it is worth noting some negative responses of the participants in introducing DDL to Thai students. To successfully investigate concordance lines and then draw generalizations on grammar rules, both students and teachers need to devote plenty of time to practice working with corpora until they become familiar enough to conduct corpus self-study. Furthermore, according to some participants, the feasibility of DDL implementation in schools is questionable as a consequence of inadequate ICT budgets or school policy proscribing in-class digital device use.

It is important to note that this study is not without limitations. One major limitation deals with the small number of participants (N = 12), which can make the use of inferential statistics less valid; therefore, only qualitative data from the interview was gathered. It is hoped that future studies will include more participants so that the results will be more generalizable. As regards the second limitation, the current study examined the perceptions of secondary teachers and tertiary educators. A clearer picture of the effectiveness of DDL in ELT will be available if DDL is applied to other study levels, e.g., primary or kindergarten. Aside from English conditionals, researchers may want to implement DDL in teaching other grammar topics. Further studies should also concentrate on the application of DDL to teaching other aspects of English outside of grammar, e.g., lexis, pragmatics, discourse, etc. It is apparent that although the efficacy of DDL in ELT has been investigated in a number of educational contexts, researchers can also consider integrating DDL with other teaching techniques to maximize students' learning outcomes (Crosthwaite et al., 2021; Lin, 2019).

References

- [1] Boontam, P. & Phoocharoensil, S. (2018). Effectiveness of English preposition learning through data-driven learning (DDL). *3L: Language, Linguistics, Literature. The Southeast Asian Journal of English Language Studies*, 24(3), 125-141.
- [2] Boulton A. (2008). DDL: Reaching the parts other teaching can't reach? *Teaching and Language Corpora*, 8, 38-44.
- [3] Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37-54.
- [4] Cowan, R. (2008). *The teacher's grammar of English. A course book and reference*. Cambridge University Press.
- [5] Crawford, W. J. & Csomay, E. (2016). *Doing corpus linguistics*. Routledge.
- [6] Crosthwaite, P., Luciana, L. & Schweinberger, M. (2021). Voices from the periphery: Perceptions of Indonesian primary vs secondary pre-service teacher trainees about corpora and data-driven learning in the L2 English classroom. *Applied Corpus Linguistics*, 1.
- [7] Dankittikul, T. & Laohawiriyanon, C. (2018). Effect of paper-based concordance on Thai low proficiency English language learners' logical connector knowledge. *Veridian E-Journal, Silpakorn University (Humanities, Social Sciences and arts)*, 11(4), 61-78.
- [8] Farr, F. & McCarthy, M. J. (2002). Expressing hypothetical meaning in context: Theory versus practice in spoken interaction. Paper read at *The Teaching and Language Corpora (TALC) Annual Conference*, Bertinoro, Italy, July 27th–30th, 2002.
- [9] Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan
- [10] Friginal, E. (2018). *Corpus linguistics for English teachers*. Routledge.
- [11] Gabrielatos, C. (2019). If-conditionals and modality: Frequency patterns and theoretical explanations. *Journal of English Linguistics*, 47(4), 301-334
- [12] Gabrielatos, C. (2021). If-conditionals: Corpus-based classification and frequency distribution. *ICAME Journal*, 45, 87-124. <https://doi.org/10.2478/icame-2021-0003>
- [13] Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(2), 225-242.

- [14] Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal*, 33(2), 99-124.
- [15] Johns, T. (1991). From print out to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14-34.
- [16] Jones, C. & Waller, D. (2015). *Corpus linguistics for grammar*. Routledge.
- [17] Karras, J. N. (2016). The effects of data-driven learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL*, 28(2), 166-186.
- [18] Koosha, M. & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL journal*, 8(4), 192-209.
- [19] Kumpawan, P. & Nishigaki, C. (2020). Active learning in the communicative English classroom using data-driven learning in Thailand. In S. Nomnian (ed). *THAITESOL Conference proceedings. Harmony in diversity: ELT in transcultural society* (pp. 122-137). Thailand TESOL Association.
- [20] Larsen-Freeman, D. & Celce-Murcia, M. (2015). *The grammar book. Form, meaning, and use for English language teachers*. The third edition. Heinle Cengage Learning.
- [21] Larsen-Walker, M. (2017). Can Data Driven Learning address L2 writers' habitual errors with English linking adverbials?. *System*, 69, 26-37.
- [22] Lastres-López, C. (2020). Beyond conditionality: On the pragmaticalization of interpersonal if-constructions in English conversation. *Journal of Pragmatics*, 157, 68-83.
- [23] Lin, M. H. (2019). Becoming a DDL teacher in English grammar classes: A pilot study. *The Journal of Language Teaching and Learning*, 9(1), 70-82.
- [24] Lin, M. H. & Lee, J. Y. (2019). Pedagogical suitability of data-driven learning in EFL grammar classes: A case study of Taiwanese students. *Language Teaching Research*, 23(5), 541-561.
- [25] O'keeffe A., McCarthy M., & Carter R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press
- [26] Petcharinphan, R. (2020). *The effects of the integrated inductive approach on GCR task and DDL in enhancing Thai EFL learners' logical connector knowledge*. [Master Thesis, Mahasarakham University]. <http://202.28.34.124/dspace/handle/123456789/1245>
- [27] Phoocharoensil, S. (2014). If-conditionals in authentic corpus-based English. *Review of European Studies*, 6(3), 62-73.
- [28] Rapti, N. (2010). *A study of classroom concordancing in the Greek context: Data-driven grammar teaching and adolescent EFL learners*. [PhD thesis]. University of Nottingham.
- [29] Rezaee, A. A., Marefat, H., & Saeedakhtar, A. (2015). Symmetrical and asymmetrical scaffolding of L2 collocations in the context of concordancing. *Computer Assisted Language Learning*, 28(6), 532-549.
- [30] Römer, U. (2011). Corpus research applications in second language teaching. *Annual review of applied linguistics*, 31, 205-225.
- [31] Sripicharn, P. (2002). Introducing data-driven learning: Theory, practice, and evaluation. *Journal of Liberal Arts, Thammasat University*, 2(2), 100-111.
- [32] Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Routledge.
- [33] Timmis, I. (2015). *Corpus linguistics for ELT*. Routledge.
- [34] Yaemtui, W. & Phoocharoensil, S. (2019). Effectiveness of data-driven learning on enhancing high-proficiency and low-proficiency Thai EFL undergraduate students' collocational knowledge. *Asian EFL Journal*, 23(3.2), 290-314.