Teacher Scaffolding in Data-Driven Learning: An Analysis of Elicitations to Assist Concordance Reading

Kunrada Chiranorawanit
Passapong Sripicharn
Thammasat University, Thailand
Corresponding author's email: kunrada.c@arts.tu.ac.th
Received August 16, 2024; revised December 29, 2024; accepted December 29, 2024

Abstract

This study investigates teacher scaffolding in Data-Driven Learning (DDL), focusing on elicitations that aid students' use of concordances for language learning. Addressing a theory-practice gap in the literature, it explores how specific teacher interactions facilitate inductive reasoning and linguistic generalizations in an EFL context. The analysis draws on classroom interaction data collected from three Thai EFL students during DDL lessons. We coded teacher elicitations using Systemic Functional Grammar's Transitivity framework, categorizing the data into material, mental, and relational clause types. The clause data was organized into concordance lines for generalizations based on the data-driven principle. The patterns found in relational and mental clauses that were used more frequently by the teacher revealed underlying efforts to engage learners in knowledge induction. Those in material clauses constructively helped learners notice and reason with tendencies in concordances. Through the collective enquiries, the teacher effectively fostered linguistic discovery and rule generalization, despite occasional counterproductive scaffolding due to over-reliance on grammatical terms. The findings highlight the importance of intentional teacher-talk strategies in DDL to support cognitive processing. Future research is invited to explore these dynamics across different linguistic topics and classroom settings.

Keywords: Data-Driven Learning (DDL), Teacher Scaffolding, Concordance Reading, Systemic Functional Grammar (SFG), EFL Classroom Interaction

Data-driven learning (DDL) promotes active work from language learners as it challenges them to be language researchers (Johns, 1991b). It incorporates big language data that is arranged in a way that increases noticeability of the target language through keyword in context (KWIC) or concordance lines. However, the transferability of knowledge from DDL comes with an instructional challenge. In the same way that language is not taught by direct telling, procedural knowledge cannot be taught with excessive teacher intervention. Hence, corpus data needs to be preserved as a language resource for the knowledge to be mined and shaped. Moreover, to facilitate learning activities, the teacher should exercise creativity in the production of DDL class materials, for example, using handouts (Lin, 2021) or a corpus interface (Nishigaki et al., 2022).

To draw the learner into the lesson, DDL has been adopted in EFL classes with a focus on specific target language features. These classes are aimed primarily at a higher level of accuracy, as measured by a pre-test and post-test. Toward this goal, concordances are pre-selected and presented on a task-sheet with questions to guide the learner toward target parts of the concordance lines, such as those aimed at vocabulary acquisition (Barabadi & Khajavi, 2017; Choksuansup & Tangpijaikul, 2017) or grammatical acquisition (Lin, 2021). Moreover, to make tasks more manageable, instructors often build a corpus with a limited number of tokens of the target item (Saeedakhtar et al., 2020), or they create an activity that encourages learners to search for answer (Curado Fuentes, 2017). As a result, learners are likely to have positive attitudes toward the tasks and to achieve the lesson goal of understanding the language use from reading concordances.

However, DDL implementation remains a challenge, and while early attempts have been made to train teachers in courses and workshops (e.g., Chen & Flowerdew, 2018; Lenko-Szymanska, 2014), it was perhaps not until Ma et al.'s (2021) two-step training—consisting of corpus literacy and corpus-based pedagogy phases—that we learned how newly introduced teachers to DDL translated their corpus linguistics knowledge into practice, and what loopholes existed in their lesson planning. The resulting corpus-based language pedagogy (CBLP), as coined by Ma et al. (2021), underscored how a teacher should use an appropriate pair of key terms and apply a required template for the

lesson procedure. Without these two things, lesson planning might not witness the same positive outcomes.

Other studies have also highlighted the difficulties of implementing DDL in various contexts. The focus on genre in Crosthwaite et al. (2021), for example, and communicative skills in Chung et al. (2024), likely made it difficult for the teacher trainees to work DDL into their lessons. The low ratings on their lesson plans also resulted from mistakes in the lesson procedure: not having a clear purpose for corpus use (lacking connection between content and pedagogy) and not specifically suggesting what the learner should find in the data (lacking instructional strategies). These oversights to put DDL knowledge into practice raise awareness as to the importance of scaffolding.

One key aspect of scaffolding is teacher talk, which is a deciding factor of not only learner satisfaction, but also the teacher's. Moreover, unreadiness to interact appears to be one main cause of anxiety in classrooms (Lin, 2019; Poole, 2022). However, the way that many teachers use concordances in the classroom may not follow the key principle of DDL, which entails that learners are supported in their inductive reasoning through teacher guidance, as encouraged by the original methodology (Johns, 1991b). As shown in Jarvis and Bayoli's (2020) study, scaffolding is key to transforming experience into deep learning. It is therefore important that the DDL teacher's advocacy of constructivism (O'Keeffe, 2021) be translated into practice. To do so, preparation of the "micro" aspect of scaffolding shall be necessary alongside that of "macro" scaffolding or the big-picture class procedure and material planning (Hammond & Gibbons, 2005).

Accordingly, our study looks into characteristics of micro scaffolding during DDL lessons. The lessons with a grammatical target were tailored to an inductive learning goal using concordances as main materials. The concordances were presented in two forms, on the computer and on paper, in two respective phases. While the use of the instructional material in the first phase led to the construction of the paper-based material in the second phase, the teacher guidance was not affected by past action. It simply grew more extensive with the reduction of the number of learners in the second sample. The interactions that went on then could be analyzed in detail to answer the question, how do teacher elicitations, organized via keywords in

context using Transitivity processes, scaffold EFL learners' inductive reasoning and linguistic generalizations during concordance-based DDL activities?

Literature Review

Concordance Reading

Although various methods may be used for concordance reading, they generally involve "rapid scanning and comparison" down the data lines (Johns, 1991b, p. 2), a process which can be guided by the Identify-Classify-Generalize procedure of Tim Johns (1991b). Within this procedure, Sinclair (2003) advises the learner to start noticing at and near the key term before working their way to the more distant context. Once some differences and similarities in the concordance lines on the page are mentally identified and classified, the learner's understanding of the given data can be generalized for other contexts inside and outside the corpus. His advice is broken down into steps as follows. First is the *Initiate* step, in which learners look at instances of a key term and its immediate co-text. Learners may notice strong patterns of a particular word form or part of speech that stand out in the data. Second, is the *Interpret* step, where learners try to induce meaning and hypothesize from the shared meaning of the repetitions. The third step is *Consolidate*, which is an interpretation that looks farther from the center for the less obvious pattern. This wider view can help learners pick up similarities and differences among larger units of meaning. Last is *Report* the finding by making a generalizing statement about the language use. Through the use of Sinclair's (2003) procedure, the learner interacts with both the language data and oneself (Hunston, 2010); hence, it is not unusual that concordance understandings differ across learners. In this study, this approach will be used to analyze the data, the result of which, albeit grounded in the data, will be open to discussion.

Classroom Concordance Reading and Learner Problems

Despite the benefits of concordances, they can be wearisome for learners who do not know basic sentence structure or have limited vocabulary. The teacher participants in Schaeffer-Lacroix (2019), for example,

believed it would be difficult for school learners to segment the text. Such claims were evidenced with the non-English major university students whose proficiency was not above school-level in Eak-in (2015); these students struggled with noticing collocations in concordances. Indeed, the DDL performance of low-proficiency and high-proficiency learners varies and is attested by the studies of Kennedy and Miceli (2010) and Yaemtui and Phoocharoensil (2019)—the less proficient learners tended not to be able to predict patterns from the data repetitions, and only learned about collocates from individual lines and picking out "good" expressions to use. Therefore, the processes of noticing patterns and interpreting shared meanings in Sinclair (2003) cannot be trusted to be autonomous for all learners.

With the elimination of gaps between learners' proficiency levels and language data levels, concordance task can still be unfulfilled when the learner refuses to take it to a higher cognitive level. This can be explained by Kim (2020), whose study found learner success and failure in undertaking the reiterative process of DDL. Looking at the data multiple times, the learners had opportunities to not merely glance at it but decipher the information and revise their ideas. This further step was blocked when they made a conclusion too soon, missing out consolidation (Sinclair's third step, 2003) and not completing "the chain of inductive reasoning" (Johns, 1991a, p. 33). As a result, their knowledge reports (Sinclair's fourth step, 2003) were not final. In other words, the learners overgeneralized their findings. This problem was found more often with the high-proficiency learners in Kim (2020) and Sripicharn (2004). More advanced in their language understanding, these learners seemed to rely on their previous knowledge, which they were already comfortable with, thus their utilization of their corpora were different from the low-proficiency learners, who could be seen to be taking time to generalize from the data.

Though learners at any level have individual problems in doing concordance tasks, they can all try and learn something from it. For instance, given an appropriate set of concordances, Sripicharn (2003) found that learners at different proficiency levels can make interesting

observations. Concordance reading is achievable with a certain language sensitivity, not necessarily erudition. In Sealy and Thompson (2004), the lower-level learners could make an effective use of concordances. They noticed "the dull words" that did not have a clearly distinguished function, construing the concept of the article. The findings from this laboratory setting nudge us to look further into classroom interactions, so that besides evidence of learning, we may also discover the kind of elicitation that supports appropriate data generalization.

Interactional Scaffolding in DDL

Concordances can be viewed as a problem set that needs solutions. The solutions are not expected to be given to the problem solver but hinted at, so that the problem is more manageable for them. Of all the things that learners can do with concordances, they are usually told to read them vertically before being asked to look for patterns. Reducing the degrees of freedom and maintaining the direction of activity as such is part of the scaffolding process according to Wood et al. (1976), who conducted a study on children doing a woodblock stacking task. Like the wooden blocks, DDL would do well with someone giving a hand to the learner to correct mistakes, model the task solution, and give purpose to the task.

Consequently, teacher interaction has been acknowledged as a driving force in DDL, but it has not been addressed as something with a recognizable pattern. Nevertheless, we should not confuse classroom discourse with lesson procedure, which for DDL was given in a formula by Tim Johns (1991b): Identify–Classify–Generalize. Urging the teacher to rely on the learners' intelligence, Johns was, however, implying that the learners would be guided by a different discourse from the typical classroom exchange that exhibits the three moves of Initiation, Response, and Feedback (IRF; Sinclair & Coulthard, 1975). But besides his encouragement for the teacher to ask the learners stimulating questions and to give them autonomy, he did not appear to give advice as to how the teacher might participate in the class conversation in a different way from the IRF pattern. Therefore, it is worth asking teachers if they

are not to use the IRF interactional pattern, how would they talk to their learners.

Interaction was later considered to be a step in the class procedure, answering the question of "when" rather than "how" to talk to learners. In the model of Carter and McCarthy (1995), Illustration-Interaction-Induction, interaction meant only peer interaction. In this model, after the learners observe concordances, they talk among themselves, then make a rule about the language item. Teacher interaction was afterwards added to this model by Flowerdew (2009), who changed the 3Is into 4Is: Illustration-Interaction-Intervention-Induction, where teacher interaction was deemed as instrumental as the other stages.

Besides these procedural formulas, there were other models with embedded teacher intervention for local applications. Nishigaki et al. (2014; cf. Brunson, 2019), for example, proposed a lesson procedure that incorporated DDL into the traditional Presentation-Practice-Production (PPP) model. After introducing a grammar point in a communicative activity in the presentation stage, a practice exercise was given to check learner understanding and whether they had any problems. This was followed be a productive task. With the addition of concordance activity, the procedure changed to Presentation–Comprehension–Practice–Production. The comprehension stage comprised the sub-stages: Look, Share, and Understand, which guided the learner to first observe the concordances, then discuss and summarize findings. The last step would be when the teacher made sure that the learners understood the grammar rule as stated in the textbook. Put another way, the DDL part of the lesson instructed for individual work, group work, and whole-class summary. Kumpawan and Nishigaki (2020) switched around the stages of individual work and group work so that the group work would build learner confidence for individual exploration. In these later models the teacher had to negotiate the learners' findings with the textbook norm.

In addition to intervention in the middle and at the end of the lesson, it is generally the case that a prompt is given at the beginning

for interactional scaffolding. This technique could spurn the learners to analyze the data, figuring out patterns of similarities and differences in them. Teacher prompts can in fact deepen the learner's understanding of the target language including vocabulary, grammar, connotation, literal/metaphorical use, and near synonyms, as shown in Table 1, as possible areas of concordance investigation (Sripicharn, 2010). Some of the prompt questions here are commonly used by teachers to guide DDL. Flowerdew (2009) suggests that this kind of intervention is especially helpful for the learners to induce phraseological patterns. She points out that the tendencies (or what learners may discover) in these patterns are harder to see than those with grammatical key terms.

Table 1

Areas of Investigation and Sample Questions (Sripicharn, 2010, p. 377)

Areas of investigation	**Sample questions**
Grammatical questions	Is 'advice' a count or non-count noun? What is the noun suffix of 'employ'?
Collocation/ phraseological questions	What are common adjectives used in front of 'argument'? What words can we use to describe our body parts (e.g. 'hair', 'nose', 'eyes')?
Connotation/ semantic prosody	What are the differences between 'childish' and 'childlike'? What are the differences between 'utterly', 'bitterly' and 'absolutely'?
Synonyms/ near synonyms	What are the differences between words in the following pairs: trip–journey; hide–conceal; attempt–try; talk–converse?
Literal/ metaphorical use	What are metaphorical uses of words literally used to describe nature such as 'flood', 'sea', 'light', 'dark'?

Even though these questions also focus the learner on specific information and can be answered by what exists in the concordance text, some questions can be more difficult than others. For instance, the key term "jerk" (Johns et al., 2008) and the key term "decreased" (Flowerdew, 2009) were used with the following prompts (see Table 2).

Table 2
Key Terms and Sample Questions from Literature

Key term	Teacher prompt
jerk	Where/what is jerked? Which instances of jerk are stronger or weaker?
decreased	Do you notice any difference in the thing that 'was decreased' and the thing that "has decreased"?
that	What feature do the citations have in common?
view	Is "view" used as a noun or a verb?

To answer the first two prompts, the learner would need to be able to, first, indicate the word that functioned as subject or object of the verb and, second, envision the relationship between the entity and the action. Unlike the open-endedness of these prompts that made the tasks divergent, the prompts for the last two key terms asked for more definite answers: "should", which is repeated in the right-hand context of the key term "that" (Johns, 1991a), and either "noun" or "verb" as regard to the key term "view" (Ackerley, 2017).

The use of prompts in DDL, which is a way to avoid knowledge prescription, requires an instructor to closely monitor the lesson. Constant vigilance would naturally result in intervention during the course of the activity, or in the middle stage of class procedure (Flowerdew, 2009). For that, there is still a lack of studies into teacher elicitations and responses in real time–only acknowledgement of learner problems such as getting hung up with cut-off sentences (Geluso & Yamaguchi, 2014), not being able to tell non-distinct patterns (Curado Feuntes, 2017), or not knowing the meaning of a word and thus learners have to read concordance lines horizontally one-by-one for the context (Yaemtui & Phoocharoensil, 2019). One of the reasons that the teacher discourse has rarely been studied in DDL could be the objective of DDL to "cut out the middleman" (Johns, 1991a, p. 30) or to maximize the chance that learner knowledge is self-constructed from looking at real language use rather than given by other people. With this ideology, the cognitive processes of individual learners

might be prioritized, with resulting negligence of social process that nonetheless have educational significance. The present study therefore taps into the social contribution of the DDL facilitator to concordance-based learning.

Method

Participants and Sample

Participants were three 11th graders, enrolled in a secondary school in Bangkok, Thailand. They were studying English as a foreign language. Their average English proficiency level was pre-intermediate. The first learner, Anna, 17, was in a study program that concentrated on English and Chinese languages. Her proficiency level was A1 Beginner. Despite the test result, she was not afraid of using English and could be considered a communicative learner. Belle, our second learner, 17, was in a study program that concentrated on mathematics and the English subjects. Her proficiency level was B1. The final learner was Chad, 17, who studied with a concentration in English and Chinese languages like Anna. His proficiency level was C1, showing excellence in the receptive skills.

The participants' teacher was the first author, who had experience with the DDL approach. She was also an in-service teacher with two years' experience and was acquainted with the school and the learners. While the other teachers in the school's English department also had rapport with the learners, they did not have any experience with the DDL approach. The teacher invited the learners to voluntarily participate in the study which was accepted with signed parental consents on the understanding that the learners could withdraw from it at any time and would bear no adverse effects, especially on their grades.

Data Collection

Data collection was done in two phases. The initial phase began with regular DDL implementation throughout the semester with one intact class of 19 students. The course was conducted online due to the COVID-19 pandemic. The second phase was conducted in the following

semester. This phase involved the three chosen participants who had received DDL lessons in the previous semester. It was carried out in a face-to-face environment. In both phases, we focused on collecting data around one teaching point, which centered on the relative pronoun "which" and the relative adverb "where", which were both part of relative clause lessons.

Phase One

A web-based corpus interface was used during Phase One. This tool was built by the first author to be user-friendly and serve a pedagogical purpose. With single-word searches on this website interface (www.kwiclook.com), learners could retrieve concordance lines from a pedagogical corpus (see Figure 1). This corpus contained approximately 100,000 words of reading texts from published ELT coursebooks, which were mostly at intermediate (B1–B2) level, reflecting the average level of proficiency of students in the class year. The language-focus lesson began with the teacher's search for a key term on a shared screen. The teacher then asked the learners to look for frequent word occurrences close to the right or left of the key term, then to notice a larger pattern and compare that to what showed up in the search results for the other key term. To help with noticing, the teacher used the cursor to mark parts of specific lines and asked questions to the whole class and random students. The DDL activity lasted about forty minutes.

Data in Phase One was collected by video recording. Recording online lessons was normal practice during COVID-19, thus it did not disturb the learners in the DDL and non-DDL parts of the English lesson.

Figure 1
Concordances on the Kwiclook Corpus Interface

Extract 1
Concordance Worksheet with the Key Terms "where" and "which"

1	Spiders are a similar color to the place	where they live.
2	I bought some pieces of metal from the shop	where my husband worked.
3	Rosie also cheated in the New York marathon	where she took the subway!
4	She got to meet Bruce Taylor, manager of the farm	where she worked as a holiday volunteer.
5	They rent rooms with sofas	where gamers can lie down after long sessions.
6	The city is full of food stands and cafés	where people meet and socialize.
7	I settled in a village not far from moscow	where I can feel at home.
8	Some bears will be released in other European countries	where they are now extinct.
9	He said, "it's a world	where everything moves fast and changes all the time."
10	He worked in his study,	where he wrote 250,000 words about his theory.
11	We've seen plenty of cookery competitions	where amateur chefs compete.
12	Here , you can see the beautiful river,	where people go sailing, and the fields beyond.
13	I take a ferry and then a bus to my school,	which is near the city centre.
14	They were swimming in an outdoor swimming pool	which was very close to the sea.
15	The tallest is one world trade center,	which is 541 metres tall.
16	A new business has grown up in Tokyo	which allows friendless people to rent a friend.
17	This is an award-winning tree house,	which a group of boys spent three years building.
18	In fact , there are boutiques	which sell a rather special line in men's clothes.
19	His son has been saved in New York,	which has been buried under a frozen sea of ice.
20	He has run the school news channel,	which he also created, and has written a series TV.
21	I'd never eaten in a place	which had table service.
22	The city is 30 minutes away from Copenhagen,	which has several hundred shops, and many brands.

Phase Two

In Phase Two, 22 concordance lines were selected from the Kwiclook corpus interface and presented on paper (see Extract 1). At the bottom were five sets of instructions. The first two parts, "Read vertically" and "Interact with the data", guided the learners to notice instances of the key term and formal patterns in the concordances. It did so by asking them to draw symbols or lines around the parts. The third part, "Focus on a few lines", drew attention to the target concordance lines. In this part, it was hoped that the learner might start to think about meaning relations between the key term and its context. Then, the last two parts, "Formulate rules" and "Fill in the blank", asked them to report what they learnt (i.e., inductive reasoning), and to check if they could apply it in novel contexts. During these final steps, the learners occasionally used iPads, as encouraged by the teacher, to check with the whole corpus data on the Kwiclook interface.

Data in Phase Two was collected through unstructured individual interviews that were conducted twice with each learner outside the school. Each session lasted about one hour. The interview started with a worksheet task, where each learner was minimally instructed. They were allowed to ask questions at any time during the task. Any unfinished tasks were to be completed in the second meeting, giving the teacher more opportunity to probe any issues.

Data Analysis

There were three main steps in data preparation. First, the data was transcribed. At this stage, data was made up of a language-related episode (LRE) from Phase One and three LREs each from the second meeting with each learner in phase two. Each LRE concerned the target language feature under focus. Second, the data was narrowed down to include only the LREs from Phase Two and to the acts of teacher elicitation, which is defined as a speech act that elicits a response (Sinclair & Coulthard, 1975). Third, the clauses of teacher elicitation were lined up in the keyword-in-context format.

In addition to simple sentences, we also included subordinate clauses, such as conditionals, if they were attached to the main clauses that acted to elicit learner responses. We also included a number of elicitative clauses in the declarative form, such as ones with a reported question embedded, "I don't know if you translate into Thai", with a question tag, "They are all used with noun, aren't they?", and imperatives that could act as both elicitation and directive, "Let's look at the other sentences whether it is not clear as well". In identifying elicitative acts, we considered the utterance meaning and tone in both the English translation and the Thai language, which was the participants' first language.

Analytical Framework

For our analysis, we drew on the view of language espoused in Systemic Functional Grammar (SFG; Halliday & Matthiessen, 2014). Within SFG, language is typically explored in terms of rank and function. Rank being the unit of analysis, and function relating to four strands of meaning that can be present in any instance of language use (experiential, logical, interpersonal, and textual functions). At the intersection of rank and function, SFG outlines system networks that represent choices language users can make (see Bowen, 2019). In this study, we focus on the clause level (rank) and experiential meanings (function), thus our analysis is based on the system network of Transitivity.

Within the Transitivity system, the clause is categorized into one of six types according to the function of the main verb, or *process type* to use SFG terminology. In this study, we adopt Halliday's view of SFG, which includes six process types: Material, mental, relational, behavioral, verbal, and existential. Associated with each of these process types are a set of attendant participants (complements in traditional grammar). Each process type has a different set of participant labels, which are outlined in detail by Halliday and Matthiessen (2014).

When change happens across space and time to the clause participant(s), the dynamic process that is transformative or creative is called *material*, as exemplified in Example 1.

Example 1

Why don't	we	use	'where'?
-	**Actor**	**Process: material**	**Goal**

Moscow	is not done to	, right?
Goal	**Process: material**	-

When the clause conveys a state or relationship between two participants, the process is called *relational*. Under the relational classification, there are sub-types including *identifying* and *attributive*, whose clause participants are reversible and irreversible respectively, as shown in Example 2.

Example 2

Is	the front part	what is done to?
Process: relational	**Identified / Identifier**	**Identifier / Identified**

'Where'	still seems	broad	, right?
Carrier	**Process: relational**	**Attribute**	-

If the emphasis of the main verb is on human sensing, the process we use is *mental*. This process represents connections between the inner and outer worlds, for which the participants are *senser* and *phenomenon*. Under the umbrella of sensing, there are several sub-processes like cognitive, desiderative, and perceptive, as realized, for example, by the verbs "understand", "want", and "see", respectively. Example 3 illustrates some mental processes as they were realized in our dataset.

Example 3

You	didn't get stuck	, right?
Senser	**Process: mental**	-

What	do	you	want to add?
Phenomenon	-	**Senser**	**Process: mental**

When	do	we	see	'which' following noun?
-	-	**Senser**	**Process: mental**	**Phenomenon**

Due to the qualitative nature of our analysis, we sought confirmation of our Transitivity coding of the Phase Two clause data from an expert. The result was 80% agreement, which was considered high, with the discrepancy lying in the borderline cases that, according to the expert's comments, were behavioral—a clause type that borders between material and mental. As we had put these clauses into either material or mental clause type, their categorizations were close to the expert's judgment, so they remained unchanged.

Results

From the Transitivity analysis, we found 12 material clauses, 29 mental clauses, and 30 relational clauses, totaling 71 clauses. The clause data were read with the process realizations in the key term positions in three concordance sets for the three Transitivity clause types. Unlike the concordance for learners above, which is made up of text from a pedagogical corpus (Extract1), the concordances below accommodate our research, with each line displaying a clause of teacher utterance to elicit learner responses (Extracts 2, 5, and 7).

Extract 2

Concordances for the Material Clauses

What did he	create?
What did she	give?
What did they	tear?
Why don't we	use 'where'?
Just now we	used 'which', right? "which he created"
Then (should) we	use 'Where did they clean up?' or…?
What did they	build?
Why does this one	use 'where' but that one 'which'?
What did the group of boys	build?
Moscow	is not done to, right?
But why does this one (line one)	use 'which'?
What did he	create?

Often in these material clauses, the actor was a human being and the goal was unknown, as seen in the empty space in the immediate right of the process. Various human actors such as "he", "she", and "a group of boys" were in the pedagogical text, whereas the goal was asked for by the question word "what", as in "What did he create?". The teacher's questions looked like those used for reading comprehension. In support of vertical concordance reading, they drew attention to certain text locations. If the learner could answer these questions, they would notice that in the target concordance lines being discussed (e.g. lines 17 and 20, Extract 1) the relative clause came after an object. These elicitations could kickstart learner observation and interpretation, the first two stages of the cognitive processing of Sinclair (2003).

In the following dialogue, Anna was guided by the teacher to have a closer look at the concordance lines where the key term, "which", was followed by a noun. She was asked via material processes, after which she expressively understood that the noun that came before "which" could also be a place—something she had thought was reserved for "where".

Extract 3

An Exchange between the Teacher and Anna

A-7	T	… 'a tree house which the boys built'…meaning… What did they build?
A-8	S	A tree house
A-9	T	'A tree house' Ah…then...for 'which she', 'a party which she gave' What did she give?
A-10	S	A party
A-11	T	A party? Ah…and there is this sentence 'He has run a school news channel which he created'. What did he create?
A-12	S	A news channel
A-13	T	A news channel …so if you could answer me what the thing is that is being done to, built, given, created...something like that, right?
A-14	S	Ahh
A-15	T	Right? For the 'which' that is a relative pronoun followed by a noun, it will mean the thing that it elaborates on - the front part, 'tree house, news channel' - must be (something).… Do you have any rule in mind? When do we see 'which' following a noun?
A-16	S	A rule in mind…mostly I think 'where' is used with place and 'which' is used with thing.
A-17	T	Ah. The rule in mind (of yours). (You) thought that 'where' was used with place and 'which' was used with thing but now...
A-18	S	But now it can be used with both.

Of interest in the data (Extract 2) were also other instances where the human actor was not in the text but in real-life, as realized by the word "we", which was followed by the material process "use". In this clause pattern, the teacher and the learner were grouped together by "we", and were cast into the role of actors, even though they were not actually manipulating or producing the goal, which was the published text. In this way, chances were created for the learner to consider why this, instead of the other linguistic choices, would be right in the sentence. They might subconsciously be conversing with their inner self: "If I were to originate it, why would I use this language item in the sentence?" In other words, being an actor in the clause could create a condition for the student to virtually put themselves in the text creator's position and to rationalize the grammatical choice. With the material process of "use", the teacher also probed for reasons using a why-

question word with the actor, “this one”, referring to a sentence that was written by the textbook author. This is another way to elicit critical comparisons, helping to consolidate the learner’s knowledge and develop their generalizations.

Consolidation and reporting are the other two important stages in concordance reading (Sinclair, 2003). The way that “use” was used in the why-questions here was contrary to traditional instruction, which would ask predominantly the question “what is used” instead of “why we use it”. As the latter promoted deeper learning, the DDL teacher did not leave the learner at the initial noticing stage but walked them through the stages after. As a result, the learner could become clearer about how they had made the generalization from the data, as the following dialogue demonstrates for the function of the relative pronoun “which” after an object. This could help make sure that the learner would be able to transfer the knowledge to other situations outside the classroom.

Extract 4

An Exchange between the Teacher and Belle

B-84	T	You wrote that ‘where’ must be followed by the subject, a pronoun, or an adjective with a noun, but ‘which’ can be followed by a verb. …How do we know that…? This is also a subject, right? 'They' or ‘there’ following ‘which’ (pointing at the data).
B-85	S	We say… I forgot to write or say that…Um, the one that follows ‘which’ is a subject that…like…is put there to just elaborate it. Right?
B-86	T	The subject is put there to elaborate on …
B-87	S	The subject is put there to elaborate on the noun.
B-88	T	Ah…(it) is put there to elaborate on it. But how?
B-89	S	For example, as I said, there is not necessarily ‘they’. It can be ‘which cleaned up this morning’ but they put ‘they’ there to let us know who the actor is - who is the one who cleans, like that. It is like something to add for explanation, right? (Referring to the sentence ‘The students could use the art room which they cleaned up this morning’)

Extract 5

Some of the Concordances for the Mental Clauses

How do you	make sense of it?
How do you	understand (it)?
Do you	understand?
You	Didn't get stuck at anywhere, right?
But I	see that you have done it sequentially, right?
So shall we	write an explanation?
I	want to ask you why you answered 'where' there
What do you	want to add?
When do we	see 'which' following noun?
I	want to know what you think.
How do we	know that…
Like…you	mean before the pronoun 'which' and 'where', right?

From general observations of the mental process concordance, the senser participants (observable to the immediate left of the process key terms) were frequently "you", referring to the learner. Following these instances of "you" were often the process realization that had something to do with comprehension such as "understand", "get", and "make sense". The teacher checked if the learners thoroughly understood the language usage based on their observation of the concordances with questions such as "Do you get it?" and "How do you understand it?"—translated from Thai (with a ring of "What do you understand?"). These questions enabled the teacher to see whether the learner had consolidated their knowledge and whether they could generalize from the data. For these purposes, the questions could also be asked during the application task, as the following dialogue shows.

Extract 6

An Exchange between the Teacher and Chad

C-1	T	Er, I want to ask you why you answered 'which' there (in a gap-filling question: 'The students could use the art room _____ they cleaned up this morning') but answered 'where' here (in a gap-filling question: 'We would like to get into a university _____ there are helpful librarians')
C-2	S	As I said at that time, the first one is clearer.
C-3	T	We would like to get into a university where there are helpful librarians.
C-4	S	This one ('where') will be too broad (for the first gap-filling item) …When we say (it) … Can you imagine it? But that one 'which they cleaned up this morning' sounds clearer. Suppose you talk to a student, they will know right away which room, right?

For the other participant sensers, namely "I" and "we" (in Extract 5), the effect could be psychological. Revealing her desire to understand by the "I" in "I want" and "I see" for instance, the teacher subconsciously portrayed herself akin to the learner who could also think and feel. Cooperating with the learner as "we", the teacher shared with the learner any difficulty and success in the task. With both "I" and "we", the equal stance could send the learner a message that their contribution was valued. It was also a conversational way to help the learner open their mind, making the task participation voluntary. So, in the last two exchange extracts above, the honest disclosure of the learners' understanding could be attributed to this kind of interrogation. This tactic of scaffolding using the mental process seemed versatile in promoting cognitive activity. It could help initiate concordance observation as in "When do we see 'which' following a noun?" as well as push for generalization as in "Shall we write an explanation?" where the phenomenon was the target of the mental process.

Although sensing processes should be commonplace in language classrooms, in DDL they supported the learners' cognitive engagement. Without the teacher's explanation of the language rule, these enquiries encouraged the learners to reflect on the linguistic phenomenon and to speak their minds based on the corpus evidence, allowing for talk of the abstract concept. Directly and indirectly, the mental processes seemingly stimulated higher-order thinking.

Extract 7

Some of the Concordances for the Relational Clauses

Ah…So in front of 'which'	is what is done to, right?
This	is also subject, right? 'they' 'there'
Ah. Yes, 'which' can	have a preposition in front, but 'where' cannot, right?
What	is active? What is passive?
This	is also 'they' (, right?)
(For) 'where',	is the front (part) what is done to?
They	are all used with noun, aren't they?
'Where'	is not used with verb, is it?
It	is 'What did they clean', right?
How	is it active, again?
Why	isn't it enough to have only 'which'?
It	will look missing something, right?

Unlike in the material and mental concordances, where humans assumed the actor and senser roles, in the relational concordance, it was more common to find the main participant to be a language item sometimes referred to as "this" or "it". Each of these preceding participants would be called the identified (item) or carrier (of an attribute). To the immediate right of the process key term was then an item or an attribute which was either a conventional grammatical term such as "subject", "preposition", "active" (voice), or an everyday description such as "the front part", "broad", or the feeling of "missing something". The fact that both grammatical and commonsensical terms were used in the teacher elicitations to discuss linguistic features suggested flexibility.

Furthermore, although we wanted to avoid deduction and traditional grammar, textbook terminology could be convenient for the teacher and the learner to communicate ideas. However, perhaps unsurprisingly, the use of textbook terminology did not always result on a positive outcome. In Extract 8, for instance, the learner, who was good at memorizing grammar, immediately picked up on a term that the teacher had used; one that bordered between commonsensical and grammatical, "what is done to", and integrated it into the concept of active and passive voice which she had studied before (turn B-101). This confusion was a helpless mistake, as the teacher did not know how much her wording was suggestive of a grammatical concept.

However, regardless of the learner factor, and whether grammatical words were used or not, the use of relational clauses could be relatively limiting. We can contrast the guidance in Extract 8 to the one in Extract 3 above and see the different effect of elicitation in the relational process (turns B-90 to B-93, Extract 8) from that of elicitation in the material process (turns A-9 to A-12, Extract 3). The building up of the learner generalization seemed more directed in one but more mysterious in the other.

Extract 8

Exchange between the Teacher and Belle (Continuing from Extract 4)

B-90	T	Ah…So in front of 'which' is what is done to, right? Since this one (following 'which') is the doer, this one (in front of 'which') must be what is done to.
B-91	S	Yes...Eh? Hold on.
B-92	T	It is what is done to like 'What did they build?' Oh, 'they built a tree house.'
B-93	S	Yes. This one is done to.
B-94	T	Ah, 'What did he create?' A school news channel, right?
B-95	S	And?
B-96	T	Therefore, it will be the subject that… (For) 'where', is the front (part) what is done to?
B-97	S	'Where'
B-98	T	'Moscow where I can feel at home' (read from the concordance line)
B-99	S	No
B-100	T	It's not. Moscow is not done to, right? …So, shall we write an explanation?
B-101	S	We write that… What? … "'Where' is not done to, but this one is done to. 'where' is passive; this one is active", right?

Another salient feature in the relational concordances (Extract 7) was the use of question tags and the sentence ending with a high intonation on "right?". Question tags appeared frequently in the far right and served to change what started off as an affirmative into an interrogative statement. Situationally speaking, the tag questions could confirm or challenge what the learner had said. They gave the learner an opportunity to change their mind, such as by the question "'Which' can have a preposition in front but 'where' cannot, right?" or to defend themselves, such as by picking on the language use in a line and asking, "This is also 'they', right?". These questions could also help

the learner move on to formulate a useful generalization. For example, by asking the question "It*'s* like the word in front of 'which' is done to, right?", the teacher gave a clue as to the area of language that could be the target of induction. Notably, this attributive clause from the interaction data seemed more effective for induction than the identifying clause ("In front of 'which' *is* what is done to, right?"). The greater degree to which the definition was hidden in the teacher's elicitation (moving down the cline from explicit to implicit saying: an object → what is done to → a possibility that x is done to), the more degrees of freedom the learner could have in giving their responses.

However, with a tentative tone uses in a statement, as illustrated in turn A-45 in Extract 9, the learner could be prevented from associating the borderline wording with grammar; the question could still entail an answer that showed the learner's genuine understanding in a very commonsensical way (turn A-48).

Extract 9

An Exchange between the Teacher and Anna (Subsequent to Extract 3)

A-35	T	Um so 'he created the news channel' (referring to the concordance line, 'He has run a school news channel which he created')
A-36	S	I think it's... like... like the back asks…like it is asking what the front does.
A-37	T	Ah. This is one of the minority cases that have a noun following 'which'. And you noted that down yourself. You remember that 'which'...
A-38	S	'Which'…
A-39	T	'Place which' and there is a noun in the back (context) …like that er mostly place or…
A-40	S	Place or word
A-41	T	The word that is in front of 'which'…
A-42	S	(The word) is 'place' and the back will be the thing that…asks what it gives. It is about… I don't know. I don't understand myself either.
A-43	T	(Giving an encouraging laugh)
A-44	S	Like, it is doing something with the front part.
A-45	T	Like it is doing something with the front part. This means that the front part is done to, right?
A-46	S	Yes. It is done to. Suppose…this one, it's cleaning an art room. The last one said creating his own news channel.
A-47	T	Um…really? Ah. And how can it be written in conclusion? Like, ...
A-48	S	What reason (is there) to have a noun following? It's the front part is a place and the back of 'which' is a thing that they...like…are doing something with the front part.

With these Transitivity processes, the learners were interactionally guided through concordance reading according to Sinclair (2003). They could initiate, interpret, consolidate, and generalize from their observation by taking the perspective of a text producer, a conscious human being, and a student studying a language concept. For the material, mental, and relational clauses, the distinct patterns in the concordances are put together in the following table.

Table 3
Summary of Observations

Process type	**Observations**		
	Frequent occurrences to immediate *left*	Frequent occurrences to immediate *right*	Frequent occurrences in the farther context
Material	Person (e.g. we, he, they)	Language item or empty	What and Why questions
Mental	Person (e.g. you, I, we)	'it', 'to' + infinitive, a clause	Why and How questions
Relational	Language item (e.g. which, where, it)	Language item, property of language item	Question tags and "right"?

Discussion

As concordance reading can be challenging, teacher intervention by probing, hinting, and eliciting becomes essential. This is in accord with socio-cultural and constructivist theories on learning. After examining teacher-talk using the framework of Transitivity, we found that the teacher questions facilitated DDL in various ways and to varying degrees. Some eliciting questions in the material, mental, and relational clause types were used subconsciously to the benefit of engaging the learners in the inductive activity and closing distance. Our findings raise awareness of the importance of teacher questioning and suggest several constructive kinds of questions that can be used in DDL classrooms.

Some of the teacher's elicitations were obviously different from those that we normally hear in the traditional language teaching classroom. Their effectiveness in getting the learner to induce knowledge

could make the linguistic configurations of these elicitations unique to and strategic for the implementation of DDL. Firstly, the realization of the material process "use" in the why-questions could probe into the learner's reasoning, helping to consolidate their knowledge. Secondly, the mental process could invite the learner to reveal more of their impressions about the target language. Thirdly, the relational process with commonsensical instead of grammatical wording, could reduce barriers to discussing linguistic concepts. Specifically, it was useful for the learner to use the terms "front" and "back" rather than grammatical terms, such as "main clause" and "subordinate clause", which were less straight-forward. In the transcript of interactions, the learner statements that indicated discovery could attest to this. For Anna, it was a better approach to concordance reading that she discovered: observing not only one but both sides of the key term. Although Belle and Chad did not enunciate their metacognition, they made statements of discovery pertaining to the target language. Belle realized that the noun following "which" gave more detail to the sentence. Chad found that the use of "which" was more specific and the use of "where" was broader. These points of view were valuable in their own rights. Moreover, the fact that each of them that was expressed in the first interview meeting and then repeated in the follow-up one suggests that what the learner had said in the first instance reflected their genuine understandings. However, at that time, the teacher may not have expected these to be the hallmark of their learning, so she kept questioning the learners as though they had been wrong.

Moreover, there were other instances where the teacher's scaffolding was unintentionally counter-productive. This stumbling block emerged when the teacher's elicitation did not respond well to the learner's discovery after it came up. The teacher would methodically point at the concordance cases that might be conflicting to the learner's assertion. This was unlikely what the learner needed. Though the interactional challenge was theoretically right, enabling the learner to re-check their understanding, it did not always allow them to work on their generalizing statements or to clarify them. For example, it did not allow Chad to dig deeper into what he meant by "broad" and "specific". Instead of refuting the learner's idea, the teacher could have been inquisitive about it, using the learner's own terms. We may keep in

mind that no matter how unorthodox or unfound the idea might sound at first, the originality of it could suggest a learning moment.

One reason for these problems could be the lack of metalinguistic tools for the learners. Unfortunately, DDL does not have that established for collaborative learning. Hence, the teacher resorted to common grammatical terms like "subject" and "verb", which could get in the way of the student's induction. An example was when the teacher's wording "is done to" caused the learner, Belle, to associate the target language with an irrelevant topic of passive voice. Another example arose during the online lesson. After the teacher stated that the relative adverb 'where' was a "subject", one of the learners took it as a clue and stated in her generalization that 'where' was followed by a "sentence". This raises a caution that the teacher's use of rule-book terminology may transfer not only the label but the concept the comes with it. It might therefore be best to avoid using grammatical terms to give learners a better opportunity to induce the answer. However, this makes us wonder how far language teaching can go without reference to grammar, since the learner will eventually need to move up to higher-level courses and develop from the present knowledge using conventional terminologies. For this, it does not help either when the teacher used a relational clause with a question tag. The tentative tone of these kinds of statements may secure learner responses, but it does not give that firmness and consistency provided in formal education for linguistic identification and classification.

Conclusion

We conducted a data-driven analysis of teacher elicitations in scaffolding data-driven learning. The concordance-based method of researching reflected that for English learning. After being categorized into the three Transitivity processes, the teacher elicitations pushed for knowledge construction in the learners' own terms. With most elicitations in the relational and mental clause types, the teacher stimulated learner contributions rather than their attention to knowledge transferred from the teacher. The use of relational or other clause types is thus cautioned to be employed with minimal mentions of conventional grammatical terms, which can interfere with the learner's meaning-making. Moreover,

it might be a better choice to use material processes, as they proved to help with patterns of noticing and differentiation. All in all, by bridging learner cognition and concordance data, teacher elicitations could scaffold inductive reasoning and rule generalizations in friendly and transformative ways.

Despite the difficulties in the interactional implementation of DDL, this study maintains that interactional scaffolding is important. We have witnessed how teacher-talk can assist learners cognitively, through the concordance reading process, enabling learners to make discoveries and grounded interpretations. Although this study did not cover the handover stage to investigate autonomy after scaffolding, some elicitations were efficient in facilitating the learners to construct rules that would be conceptually and procedurally memorable to them.

Overall, while we undertook the case of grammar teaching, which involved a great deal of abstraction, future research may investigate teacher elicitation on other language points like vocabulary. It will be interesting to see whether they are taught with the same clause types and linguistic features. In addition to the language topic, the classroom setting can be a criterion that limits these findings to the conduct of our study. Most of our interactional data was collected from individual lessons in which the guidance was quite informal. Probing in the manner of interviewing may not be practical with many students; the lightness in the teacher-talk, which is uncommon in Thai school classes, also risks being misinterpreted as digression. To benefit a wider DDL implementation, teacher interactions in face-to-face classroom in the Thai setting should be investigated. Ultimately, considering the preference for standard practice, we need to be prepared for the unpredictable that could happen in human-to-human interactions, which seems unavoidable in the inductive approach to learning.

References

Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology*, *21*(3), 195–216.

Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning English vocabulary. *Cogent Education*, *4*(1). https://doi.org/10.1080/2331186X.2017.1283876

Bowen, N. E. J. A. (2019). Unfolding choices in digital writing: A functional perspective on the language of academic revisions. *Journal of Writing Research*, *10*(3), 465–498. https://doi.org/10.17239/jowr-2019.10.03.03

Brunson, H. S. (2019). *Developing intermediate-level ESL grammar lessons for community college students using guided data-driven induction* [Master's thesis]. California State University.

Carter, R., & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, *16*(2), 141–158. https://doi.org/10.1093/applin/16.2.141

Chen, M., & Flowerdew, J. (2018). Introducing data-driven learning to PhD students for research writing purposes: A territory-wide project in Hong Kong. *English for Specific Purposes*, *50*, 97–112. https://doi.org/10.1016/j.esp.2017.11.004

Choksuansup, N., & Tangpijaikul, M. (2017). The effect of concordancing-based teaching on vocabulary learning and retention on Thai EFL learners. *Chophayom Journal*, *28*(3), 11–19.

Chung, T. T., Crosthwaite, P., Cao, C. T. H., & De Carvalho, C. T. (2024). Walking the walk? (Mis)alignment of EFL teachers' self-reported corpus literacy skills and their competence in planning and implementing corpus-based language pedagogy. *TESOL Quarterly*, 58(1). https://doi.org/10.1002/tesq.3299

Crosthwaite, P., Luciana, & Wijaya, D. (2021). Exploring language teachers' lesson planning for corpus-based language teaching: A focus on developing TPACK for corpora and DDL. *Computer Assisted Language Learning*, 1–29. https://doi.org/10.1080/09588221.2021.1995001

Curado Fuentes, A. (2017). Form-focused data-driven learning for grammar development in ESP contexts. *Revista de Lenguas para Fines Especificos*, *23*(1), 12–30. https://doi.org/10.20420/rlfe.2017.155

Eak-in, S. (2015). Effects of a corpus-based instructional method on students' learning of abstract writing: A case study of an EAP course for engineering students. *Journal of English Studies, 10*, 1–41.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy - A critical evaluation. *International Journal of Corpus Linguistics*, *14*(3), 393–417. https://doi.org/10.1075/ijcl.14.3.05flo

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, *26*(2), 225–242. https://doi.org/10.1017/S0958344014000044

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect*, *20*(1), 6–30. https://neilwhitfield.wordpress.com/wp-content/uploads/2008/11/20_1_1_hammond.pdf

Hunston, S. (2010). How can a corpus be used to explore patterns? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed., pp. 152–166). Routledge.

Jarvis, M., & Baloyi, O. B. (2020). Scaffolding in reflective journaling: A means to develop higher order thinking skills in undergraduate learners. *International Journal of Africa Nursing Sciences*, *12*, Article 100195. https://doi.org/10.1016/j.ijans.2020.100195

Johns, T. (1991a). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Johns & P. King (Eds.), *ELR Journal*, *4*, 27–45.

Johns, T. (1991b). Should you be persuaded - Two samples of data-driven learning materials. In T. Johns & P. King (Eds.), *ELR Journal*, *4*, 1–16.

Johns, T., Hsingchin, L., & Lixun, W. (2008). Integrating corpus-based CALL programs in teaching English through children's literature. *Computer Assisted Language Learning*, *21*(5), 483–506. https://doi.org/10.1080/09588220802448006

Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology*, *14*(1), 28–44. https://doi.org/10125/44201

Kim, M. (2020). A qualitative analysis of EFL learners' discrimination of near-synonyms in a data-driven task. *English Teaching*, *75*(3), 25–47. https://doi.org/10.15858/engtea.75.3.202009.25

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* (pp. 122–137). n.p.

Lenko-szymanska, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL*, 26(2), 260–278. https://doi.org/10.1017/S095834401400010X

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.

Lin, M. H. (2021). Effects of data-driven learning on college students of different grammar proficiencies: A preliminary empirical assessment in EFL classes. *SAGE Open*, *July-September*, 1–15. https://doi.org/10.1177/21582440211029936

Ma, Q., Tang, J., & Lin, S. (2021). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, *35*(9), 1–30. https://doi.org/10.1080/09588221.2021.1895225

Nishigaki, C., Akasegawa, S., & Oghigian, K. (2022). Development of an online DDL tool for secondary school learners. *Cukurova University Faculty of Education Journal*, 70, 289–297. https://doi.org/10. 20776/S13482084-70-P289

Nishigaki, C., Kunikane, A., Yoshida, S., Sanada, N., & Chujo, K. (2014). Application and effects of data-driven learning at junior high school (in Japanese). *Bulletin of the Faculty of Education, Chiba University*, *60*, 419–426.

O'Keeffe, A. (2021). Data-driven learning - A call for a broader research gaze. *Language Teaching*, *54*, 259–272. https://doi.org/10.1017/S0261444820000245

Poole, R. (2022). "Corpus can be tricky": Revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, *35*(7), 1620–1641. https://doi.org/10.1080/09588221.2020.1825095

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*, *91*, 1–14. https://doi.org/10.1016/j.system.2020.102268

Schaeffer-Lacroix, E. (2019). Barriers to trainee teachers' corpus use. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners* (pp. 47-64). Routledge. https://doi.org/10.4324/9780429425899

Sealey, A., & Thompson, J. (2004). 'What do you call the dull words?' Primary school children using corpus-based approaches to learn about language. *English in Education*, *38*(1). https://doi.org/10.1111/j.1754-8845.2004.tb00774.x

Sinclair, J. (2003). *Reading concordances: An introduction*. Longman Pearson Education.

Sinclair, J., & Coulthard, M. (1975). Chapter 3. In *Towards an analysis of discourse* (pp. 14–60). n.p.

Sripicharn, P. (2003). Evaluating classroom concordancing: The use of concordance-based materials by a group of Thai students. *Thammasat Review*, *8*(1), 203–236.

Sripicharn, P. (2004). Examining native speakers' and learners' investigation of the same concordance data and its implications for classroom concordancing with EFL learners. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 233–245). John Benjamins Publishing. https://doi.org/10.1075/scl.17.17sri

Sripicharn, P. (2010). How can we prepare learners for using language corpora? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 371–384). Routledge.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yaemtui, W., & Phoocharoensil, S. (2019). Effectiveness of data-driven learning (DDL) on enhancing high-proficiency and low-proficiency Thai EFL undergraduate students' collocational knowledge. *Asian EFL Journal*, *23*(3), 290–314.