

Ana Mužar*

University of Belgrade
School of Dental Medicine
English Language Department

 <https://orcid.org/0009-0003-6158-7861>

Irena Aleksić-Hajduković**

University of Belgrade
School of Dental Medicine
English Language Department

 <https://orcid.org/0000-0002-1884-4784>

CONVERGING GENAI AND DATA-DRIVEN LEARNING FOR ESP MATERIALS DEVELOPMENT

Original scientific paper
UDC 004.8
811.111`276.5/.6:616.314
811.111`243
37.091.64
37:159.953]:001.103
<https://doi.org/10.18485/kkonline.2026.17.s2.2>

Data-driven learning (DDL) refers to an effective methodological approach to language teaching and learning that relies on using language corpus data and tools in combination with suitable pedagogical approaches (Crosthwaite & Boulton, in press). However, with recent developments in the field of Generative Artificial Intelligence (GenAI), there has been a need to merge the two approaches in order to bridge the existing gaps and transcend the limitations of DDL (Crosthwaite & Baisa, 2023; Mizumoto, 2023). Therefore, the current study exploits a corpus comprised of dental research abstracts generated by means of AntCorGen and analysed via Sketch Engine software tools in order to illustrate how DDL and GenAI can complement each other when it comes to assisting ESP materials design. For the purposes of designing an array of exercises and tasks intended for mixed-ability learners of English for Dental Medicine at a tertiary level, the inspiration has been drawn from *Cambridge GenAI Idea Pack* that offers teacher-friendly and easy-to-use prompt models applicable in various classroom settings. The assortment of classroom activities focused on field-specific collocation acquisition and tailored to our students' specific learning context includes multiform vocabulary and listening exercises, as well as inductive learning activities. This approach to ESP materials design offers practical implications for converging corpus data and GenAI with the aim of facilitating student exposure to authentic language use.

Keywords: Corpus Linguistics, collocations, DDL, General Artificial Intelligence, English for Healthcare, problem-based learning, role-plays.

* University of Belgrade, School of Dental Medicine, 8, Dr Subotića, 11000 Belgrade, Serbia; e-mail: ana.muzar@stomf.bg.ac.rs

** University of Belgrade, School of Dental Medicine, 8, Dr Subotića, 11000 Belgrade, Serbia; e-mail: irena.aleksic@stomf.bg.ac.rs

1. Introduction

Ever since Johns (1991) first coined the term data-driven learning (DDL), it has been used to denote an effective methodological approach to language teaching and learning that relies on using language corpus data and tools in combination with suitable pedagogical approaches (Boulton & Pérez-Paredes, 2024; Crosthwaite & Boulton, in press; Csomay, Jablonkai, & Sun, 2025).

The DDL approach is particularly informative and useful in the context of English for Specific Purposes (ESP). Namely, evidence obtained through corpora explorations provides quantitative data to guarantee the commonness or typicality of a word and can aid in differentiating the nuances of a word in a general or specialized domain of real-life communication (Xu, 2022, p. 17). This is of supreme importance in ESP as we strive to empower our learners to communicate effectively within a plethora of professional domains, each with its field-specific terminology and varying communicative needs. In terms of materials development, Xu (2022, p. 17) emphasises that corpora can contribute by providing samples of real-life language use, vocabulary control, and a phraseological approach to lexis and grammar.

Csomay et al. (2025) remind us that the degree of scaffolding and direct access to corpus data determine the prevalent approaches to corpus-based pedagogy, which include corpus-informed teaching practice, integrated corpus-supported teaching, and self-directed DDL. These authors also point out a meta-analysis recently conducted by Liu and Ma (2025) who reported that for low-proficiency L2 English learners approaching corpus data through paper-based DDL, i.e. without students' direct interaction with corpora, is more likely to yield favourable results in relation to improving lexico-grammatical knowledge. Thereby, instructional scaffolding and feedback should be ensured as key prerequisites for successful implementation of this approach. In the context of ESP, it could be argued that this translates to both designing and navigating students through corpus-informed activities with the aim of enhancing the acquisition of subject-specific terminology and communicative skills as we mostly work with students with varying levels of both productive and receptive English language skills. In summary, through regularly adapting and upgrading corpus-based materials, language educators can create prolific learning environments that mimic real-world communication (Friginal & Prado, 2025).

However, DDL is also associated with certain drawbacks that, for instance, Sun and Mizumoto (2025) have identified and described *in extenso* in their systematic qualitative synthesis of 347 primary studies published from 2012 to 2021. Their findings

suggest that the challenges linked to DDL implementation can be classified into *inherent* (pertaining to the DDL tasks integrated, e.g. time constraints, interface issues, preference for other resources, accessibility, and misalignment of purposes) and *external* factors (pertaining to DDL users, i.e. learners, teachers, and institutional support) (Sun & Mizumoto, 2025; italics added for emphasis). One of the key findings of this study is that the core of the obstacles related to DDL integration surpasses the technical level, but student capacity and beliefs they shared stress that the emotional aspect should inevitably be considered in terms of user engagement. This conclusion has a direct impact on the tools, teaching methods and design of study materials associated with the DDL approach since it emphasises emotional aspects such as motivation, confidence, likelihood for future use, learner autonomy, etc. It is important to note that the studies included in this systematic qualitative synthesis were all published before ChatGPT, which was made available in late 2022, galvanised people around the world into experimenting with the possibilities of GenAI-powered tools.

It was not long afterwards that notable scholars in the realm of DDL (Crosthwaite & Baisa, 2023; Mizumoto, 2023) put forward the idea of converging DDL and Generative Artificial Intelligence (GenAI) in an attempt to overcome the challenges and obstacles associated with DDL and corpus-informed language learning. This seems especially salient in the field of ESP as we teach undergraduate students whose primary goal is not language education, regardless of how much effort we put into illustrating real-life challenges arising from inadequate professional communication in an intercultural environment. In this paper, we will explore the potential of GenAI for ESP materials development and not the integration of GenAI into the existing corpora tools, which is another promising feature of GenAI implementation. One such example is the addition of *ChatAI* to *AntConc* corpus software (Laurence, 2025a) which, at the time of writing this paper, enables first-hand interactions with OpenAI's GPT-5 models not only when designing corpus queries, but also when exploring the outcomes of a corpus query (Laurence, 2025b).

Previous research pinpoints that AI should not be perceived as a substitute for corpora tools or data-driven approaches to language teaching and learning but rather as an additional resource that has the potential to boost the effectiveness of DDL and corpus-based pedagogy (Csomay, Jablonkai, & Sun, 2025; Pérez-Paredes & Boulton, 2025). Crosthwaite and Baisa (2023) list a number of reasons to underpin this stance, including *knowing the data, authenticity, replicability, etc.* *Knowing the data* could be reconstructed as the record of the process of corpus compilation and analyses performed

that led us to the corpus-based findings. Such a trace is non-existent when large language models (LLMs) such as ChatGPT are used (Crosthwaite and Baisa, 2023), e.g. to provide information about the keywords or collocations in an ESP domain such as English for Dental Medicine. Even though Pérez-Paredes and Boulton (2025, p. 72) state that students with no training in corpora tools can now use ChatGPT to gather information about nodes and collocates of their interest, including contextualised samples of their usage; they also point out *authenticity* as an issue previously defined by Crosthwaite and Baisa (2023) as “language data *actually produced by humans*”. Finally, by *replicability* Crosthwaite and Baisa (2023) refer to the fact that, at this point, end-users cannot see, nor replicate, the statistical procedures that resulted in the generated output.

Finally, the issues stemming from implementing AI technologies in the classroom should not be neglected either. As Edmett et al. (2023, p. 52) underline, there is no doubt that “AI literacy is still a developing teacher training area”. Understandably, many language teachers may feel like being on a slippery slope when employing GenAI tools for the purposes of creating classroom exercises, tasks, lesson plans, and such due to insufficient training opportunities encompassing both prompt-engineering practices and practical implementation of GenAI in a language learning environment (Edmett et al., 2023, p. 30).

The present study focuses on collocations in the field of dental medicine obtained by corpus-analysis methods and their incorporation into GenAI for ESP materials design. The reason for investigating the field-specific units of formulaic language used in dental medicine lies in the fact that a scarce number of studies have focused on learning different types of collocations in the context of English for Specific Purposes (ESP), which is particularly perturbing bearing in mind that for effective communication in academic and professional settings acquiring discipline-specific collocations is of utter importance (Farooqui, 2025).

Based on the rationale behind merging DDL and GenAI, in this paper we set out to explore the language of dental medicine used in dental research abstracts with the aim of designing a research-backed set of dental collocations that will subsequently be used as the foundation for designing GenAI-powered study materials in this domain of ESP. Such a deconstruction of the process of integration of DDL and GenAI is applicable to a broad spectrum of ESP domains that could benefit from experimenting with this novel pedagogical approach to language learning.

2. Methodology

2.1. Corpus compilation and analysis

The methodological approach to the corpus investigated in the present study consists of two phases – corpus compilation¹ and corpus analysis. First, in order to create a set of target collocations that would later be incorporated into ChatGPT, it was necessary to create a corpus comprising texts from the field of dental medicine (DENTRA – DENTal Research Abstracts). The freely available software AntCorGen v. 1.3.0 (Anthony, 2024) allows searching through the PLOS One database and was used to compile a corpus of dental research abstracts belonging to the category of *Oral Medicine* (subcategories: *Dentistry* and *Oral Health*), as shown in Figure 1. The specialised fields available within this domain were not included in the study since we aimed to create exercises suitable for our first-year students. The end result was the corpus comprising 419 abstracts, i.e. 109,269 tokens.

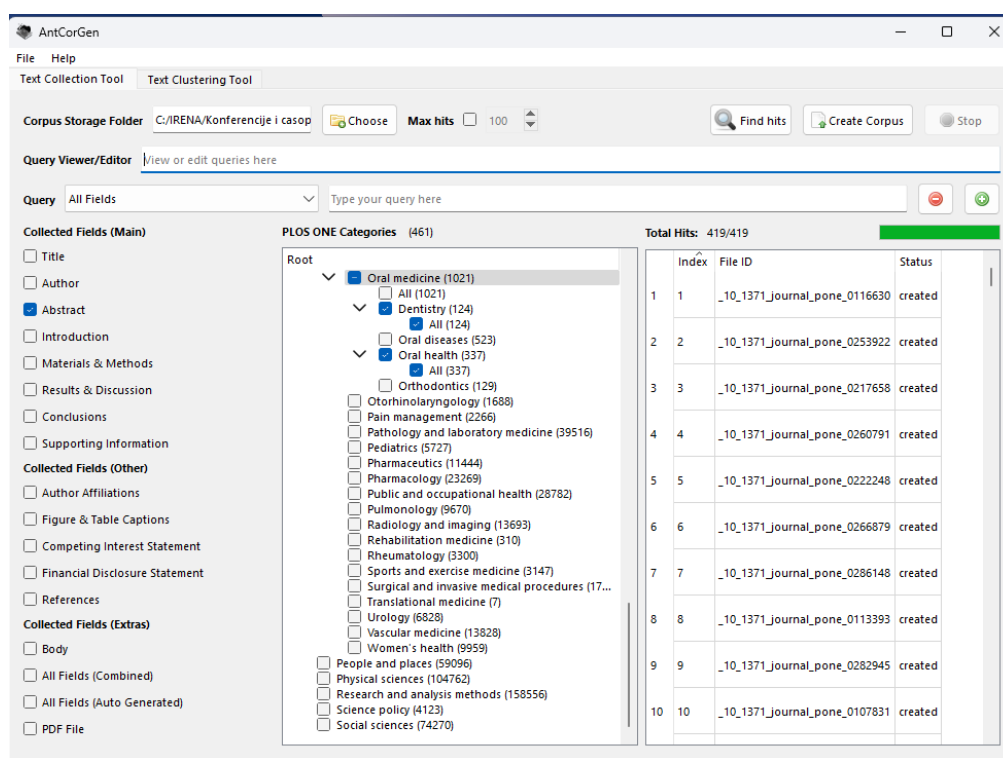


Figure 1: DENTRA corpus compilation using AntCorGen v. 1.3.0

Subsequently, Sketch Engine (Kilgariff et al., 2014) was used to conduct a keyword analysis, accompanied by the Word Sketch analysis of the selected nodes as it

¹ For the step-by-step procedure of corpus compilation using AntCorGen, see Asano (2018).

represents a corpus-based summary of a word's grammatical and collocational behaviour (Baisa, 2010, p. 37). The search span is set from -5 to 5 in Sketch Engine, i.e. it encompasses up to 5 words left from the node (-5) and up to 5 words found on the right side of the node. The *British National Corpus (BNC)*, more precisely its *Written Academic* subcorpus, was used to identify the keywords in the field of dental medicine rather than the keywords typical of the academic register. While the keyness score was calculated using the simple maths method, the typicality of a collocation was determined based on the logDice score, both representing the default settings in Sketch Engine. Section 3 contains the findings obtained through these analyses.

2.2. Prompt engineering

Technically speaking, the generation of AI content is directly contingent upon prompts produced in natural-language conversational interfaces (UNESCO, 2023, p. 11). For the purpose of formulating effective prompts, we relied on *Cambridge Generative AI Idea Pack for English Language Teachers* (2024) (hereafter *Cambridge GenAI Idea Pack*) which contains teacher-friendly, easy-to-use prompt models that can yield tangible results in the form of diverse exercises applicable in various classroom settings. What this *Cambridge GenAI Idea Pack* (2024) suggests when experimenting with GenAI tools in this respect is feeding as many necessary details as possible into an AI-based chatbot, including the target vocabulary, as well as the style and the length of the AI-generated text. Adopting such an approach to prompt formulation increases the text relevance to the target group we are teaching and helps us arrive at a highly customised GenAI output that caters to the specificities of our students' instructional setting. However, the original prompts from *Cambridge GenAI Idea Pack* are aimed at General English learners; thus, they were revised and adapted so as to focus on the set of our target collocations compiled based on the corpus data. These prompts were then incorporated into ChatGPT 4o-mini (See Section 3), which was available at the time. With regard to the prompt engineering process, mindfulness of the necessity for teacher revision and constant critical evaluation needs to be highlighted.

Besides, it should be stressed that both accuracy and relevance of all AI-generated content presented in this paper were verified by an expert in the field of dental medicine employed at our institution. This form of cooperation between language teachers and content-based subject specialists is of paramount importance especially when language teachers are not able to check for validity of the AI-generated output themselves by cross-referencing it with more reliable resources.

To sum up, one of the objectives of this paper is to showcase how merging corpus-based data and GenAI tools with the prompt-engineering guidance provided by e.g. *Cambridge Generative AI Idea Pack* (2024) can facilitate the design of ESP study materials, which will be further elucidated in the following section.

3. Results and Discussion

3.1 Creating an evidence-backed set of target collocations

In this section, we will present the main findings of the DENTRA corpus exploration in order to illustrate the process of designing an evidence-backed collocational set. Figure 2 shows the top 20 keywords in the DENTRA corpus. The ones highlighted have been selected as the nodes for the collocational analysis performed via Word Sketch.

Lemma	Frequency [?]		Frequency per million [?]		Score [?]
	Focus	Reference	Focus	Reference	
1 caries	330	9	2,383.69	0.51	1,578.6 ...
2 dental	1,071	119	7,736.15	6.75	998.1 ...
3 dentistry	112	7	809.01	0.40	579.8 ...
4 healthcare	97	7	700.66	0.40	502.2 ...
5 dmft	67	0	483.96	0.00	485.0 ...
6 dentist	224	47	1,618.02	2.67	441.6 ...
7 behavior	125	24	902.91	1.36	382.7 ...
8 health-related	59	4	426.17	0.23	348.2 ...
9 ohrqol	46	0	332.27	0.00	333.3 ...
10 periodontal	179	51	1,292.97	2.89	332.3 ...

Lemma	Frequency [?]		Frequency per million [?]		Score [?]
	Focus	Reference	Focus	Reference	
11 cross-sectional	117	30	845.13	1.70	313.1 ...
12 periodontitis	73	13	527.30	0.74	304.0 ...
13 covid-19	41	0	296.16	0.00	297.2 ...
14 ohi	40	0	288.93	0.00	289.9 ...
15 self-reported	72	16	520.08	0.91	273.1 ...
16 ecc	37	0	267.26	0.00	268.3 ...
17 gingival	37	2	267.26	0.11	240.9 ...
18 oht	32	0	231.15	0.00	232.1 ...
19 dentition	34	2	245.59	0.11	221.5 ...
20 cpq	31	1	223.92	0.06	212.8 ...

Figure 2: Top 20 keywords in the DENTRA corpus (Sketch Engine, keyness score: simple maths)

Even though corpora tools allow us to identify the keywords in a corpus, this is not to say that we should focus on top-ranked ones exclusively. Instead, we opted for those we considered most useful for our dental students based on our course syllabus and the instructors' classroom experience (*caries*, *dental*, *periodontal*, *gingival*, and *dentition*; ranking as 1, 2, 10, 17, and 19, respectively). For instance, *dentition* ranked as No. 19. Therefore, including e.g. the top-5 keywords would mean that *dentition* would be omitted. However, *human dentition* represents an important segment of our course syllabus and from the authors' hands-on experience, students find it difficult to remember which adjectives co-occur with *dentition* (permanent, mixed, inadequate, etc.).

Next, we performed a Word Sketch analysis for the five nodes and then selected five out of the most typical collocates identified by the software. The default setting classifies collocates based on typicality and not frequency. The logDice score is used for calculating how typical a collocation is. In other words, this measurement is indicative of how strong the connection between the node and a collocate is (Sketch Engine, n.d.). Table 1 contains the chosen nodes and their collocates. The number in the bracket represents the ranking of the collocate based on the logDice score.

NODE	CARIES	DENTAL	PERIODONTAL	GINGIVAL	DENTITION
COLLOCATE	experience (2)	caries (1)	disease (1)	bleeding (1)	functional (1)
	prevalence (3)	care (2)	pocket (2)	margin (2)	permanent (2)
	childhood (4)	services (3)	status (3)	inflammation (3)	inadequate (3)
	untreated (7)	treatment (4)	therapy (4)	index (4)	non-functional (6)
	advanced (8)	practitioner (12)	depth (15)	swab (6)	mixed (16)

Table 1: The collocational data set including the chosen nodes, collocates, and their rankings in the DENTRA corpus (Sketch Engine, Word Sketch, logDice score)

Eventually, a set of 25 collocations was prepared and ready to be used for designing GenAI-powered study materials. Despite the fact that these collocations have been included in the materials design presented in this paper, different datasets can be comprised by including different collocates of these keywords or other keywords and their collocates in alignment with a particular topic discussed in the class and students' specific needs.

In the pursuing sections of this paper, we will discuss the process of prompt formulation and outputs received with regard to designing vocabulary exercises, reading and listening exercises, case studies, and problem-based tasks. We will also display how GenAI can be prompted to suggest suitable ELT methodologies for the purpose of teaching target collocations.

On a related note, it is not only about what we believe our students should learn, but it is about designing meaningful and research-backed activities, while we scaffold them throughout the process. This is one of the reasons why the role of the human educator cannot be overemphasised in the AI era.

3.2 Creating ChatGPT-powered vocabulary exercises

AI-powered chatbots appear to be particularly convenient for generating field-specific texts that abound in ESP vocabulary (Kovačević, 2024; Warschauer et al., 2023) and, by extension, transforming those texts into a variety of vocabulary exercises (Generative AI Idea Pack, 2024)².

The following prompt (*Revised prompt 1*) was created by revising the *Test generation* prompt presented in *Cambridge GenAI Idea Pack* (2024). The prompt was modified to be suited to a mixed-ability class of ESP university students practising dental collocations related to gingivitis.

Revised prompt 1	Generate a 100-word text for intermediate learners of English for Healthcare at a tertiary level and include these words: gingival bleeding, gingival margin, gingival inflammation, gingival index, gingival swab. Remove the vocabulary items specified previously and substitute them with gaps. Make the gaps the same length. Provide the vocabulary items as a list to choose from under the text in a mixed order.
-------------------------	---

In order to tailor the ChatGPT output to the specific needs of our target student population, the original prompt was adapted to encompass their proficiency level and the particularities of their learning context as well. Additionally, a few critical points were noted during the prompt formulation stage and as such warrant particular emphasis. First of all, adding a letter of the alphabet in front of each gap, as stated in the original prompt, did not prove a neat solution aesthetically and was, therefore, excluded altogether from the revised prompt. Secondly, the generated gaps were not the same length and the vocabulary items provided as a list to choose from under the text were not given in a mixed order, which necessitated further refinement and iterative prompt adjustments.

Figure 3 depicts a ChatGPT-generated gap-fill exercise that contains the target collocations pertaining to the node *gingival*.

² Only the revised prompts are presented in the paper. For original prompts, please see *Cambridge Generative AI Idea Pack for English Language Teachers* (2024).
https://www.cambridge.org/sites/default/files/media/documents/GenAI_Idea_Pack.pdf

Text:

During a routine dental examination, the dentist noticed signs of _____, particularly at the _____. To assess the severity of the condition, they measured the _____, which revealed moderate levels of inflammation. The patient reported occasional _____ during brushing, which was consistent with the observed symptoms. To further investigate, the dentist took a _____ to analyze for any bacterial infections contributing to the inflammation.

Vocabulary List (in mixed order):

1. Gingival bleeding
2. Gingival index
3. Gingival margin
4. Gingival swab
5. Gingival inflammation

Figure 3: A ChatGPT-powered gap-fill exercise based on the collocates of the node *gingival*

This capability of GenAI tools to craft field-specific texts and accompanying exercises may also come in handy in instances where no official coursebook has been assigned for an ESP course (Cambridge GenAI Idea Pack, 2024).

Next, the prompt entitled *Write your own text* was adapted from *Cambridge GenAI Idea Pack* (2024) to aid a mixed-ability class of ESP university students in learning about *dentition* through collocational practice (*Revised prompt 2*).

Revised prompt 2	I’m an ESP teacher at a tertiary level. I teach intermediate students of English for Healthcare. They study English for Dental Purposes. Write an informative article of 150 words about dentition. Include the following vocabulary: functional dentition, permanent dentition, inadequate dentition, non-functional dentition, mixed dentition. Create some vocabulary exercises based on this text.
-------------------------	--

This time, the original prompt included students’ proficiency level, as well as their specific learning context. ChatGPT was additionally prompted to propose vocabulary exercises that could accompany the subject-specific text, which resulted in a range of exercise types, including gap-fill, matching, multiple choice, and T/F formats.

Figures 4-8 portray the ChatGPT-generated text and follow-up vocabulary exercises designed to help ESP students acquire collocations relative to the node *dentition*.

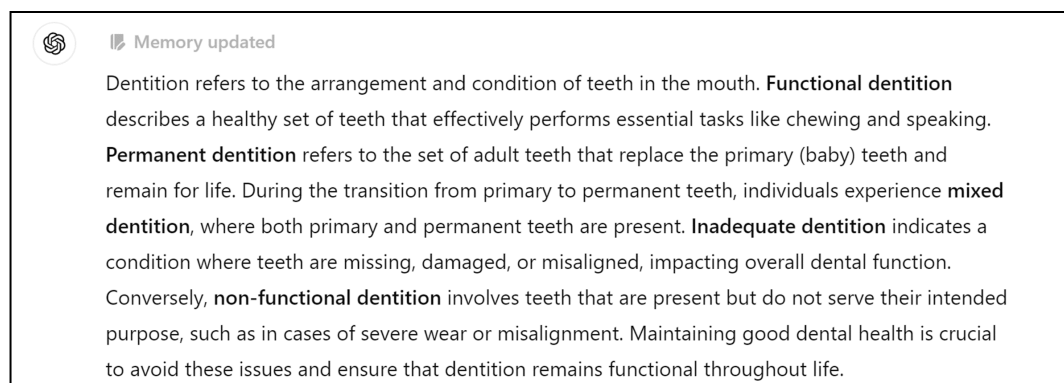


Figure 4: A ChatGPT-generated informative article containing the selected collocates of the node *dentition*

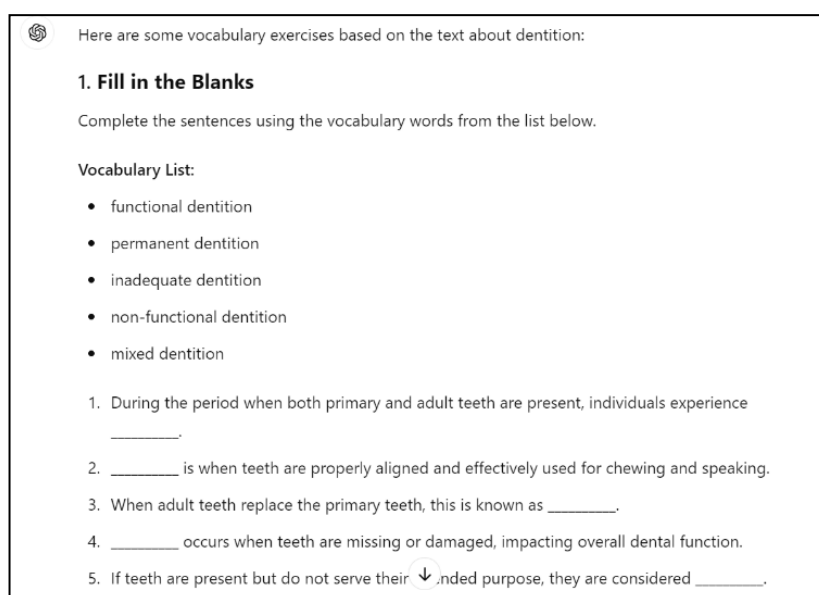


Figure 5: A ChatGPT-powered gap-fill vocabulary exercise targeting the selected collocates of the node *dentition*

2. Match the Terms

Match each term to its correct definition.

Terms:

- Functional Dentition
- Permanent Dentition
- Inadequate Dentition
- Mixed Dentition
- Non-Functional Dentition

Definitions:

1. The set of adult teeth that replace primary teeth and remain for life.
2. Teeth that are properly aligned and serve their intended functions, like chewing.
3. A condition where teeth are missing or misaligned, affecting dental function.
4. The condition when both primary and permanent teeth are present.
5. Teeth that are present but do not perform their intended functions effectively.

Figure 6: A ChatGPT-powered matching vocabulary exercise targeting the selected collocates of the node *dentition*

3. Multiple Choice

Choose the correct answer for each question.

1. What is the term for the set of adult teeth that replaces primary teeth?
 - a) Functional Dentition
 - b) Permanent Dentition
 - c) Mixed Dentition
2. Which term describes teeth that are not performing their intended functions effectively?
 - a) Non-Functional Dentition
 - b) Inadequate Dentition
 - c) Mixed Dentition
3. What is experienced during the transition from primary to adult teeth?
 - a) Permanent Dentition
 - b) Functional Dentition
 - c) Mixed Dentition
4. What term refers to a condition where teeth are missing, damaged, or misaligned?
 - a) Functional Dentition
 - b) Inadequate Dentition
 - c) Non-Functional Dentition
5. Which term indicates a healthy set of teeth that performs essential tasks like chewing?
 - a) Non-Functional Dentition
 - b) Functional Dentition
 - c) Mixed Dentition

Figure 7: A ChatGPT-powered multiple choice vocabulary exercise targeting the selected collocates of the node *dentition*

4. True or False

Determine if the statements are true or false.

1. **Functional Dentition** includes teeth that are not used for chewing.
2. **Mixed Dentition** occurs when both primary and adult teeth are present.
3. **Permanent Dentition** refers to the set of teeth that replace primary teeth and are intended to last a lifetime.
4. **Inadequate Dentition** means teeth are perfectly aligned and serve all functions.
5. **Non-Functional Dentition** involves teeth that do not serve their intended purpose, such as severely worn or misaligned teeth.

Figure 8: A ChatGPT-powered True/False vocabulary exercise targeting the selected collocates of the node *dentition*

Notably, having been engaged in creating a gap-fill exercise for our target student population, ChatGPT now displays updated memory on the screen. In other words, the learning context is inputted right away in the ChatGPT memory and all the exercises subsequently created are generated in relation to the previously inputted background information.

As for the ChatGPT-generated follow-up vocabulary exercises, it is worth mentioning that even though substantive post-editing was not required, a selective approach to choosing which exercises exactly to include in the classroom is needed due to their marked similarity in content and question formulation – or, alternatively, if opting for all four of them, length extension and further fine-tuning of the prompt are warranted.

3.3 Creating ChatGPT-powered reading and listening exercises

In a similar vein, and with the very same prompt adjusted accordingly, ChatGPT can be employed for the purposes of creating both reading and listening comprehension exercises. In fact, Edmett et al.'s (2023, pp. 20, 47) systematic literature review indicates that a stronger research focus should be placed on how AI can be implemented in an ELT classroom to foster the development of receptive skills, inasmuch as reading and listening were deemed less "ripe for transformation via AI" than speaking and writing at the time of conducting this review of the literature. Therefore, it should come as no surprise that only 19% of the English language teachers encompassed by Edmett et al.'s (2023, p. 24) study reported using text-to-speech tools as part of their AI-powered tool application in ELT.

Revised prompt 3 demonstrates how the *Write your own text* prompt was customised, once again, to address the specificities of our ESP students' learning context and assist them in practising collocations that pertain to *dental caries*:

Revised prompt 3	I'm an ESP teacher at a tertiary level. I teach intermediate students of English for Healthcare. They study English for Dental Purposes. Write an informative article of 200 words and include the following vocabulary: caries experience, the prevalence of dental caries, childhood caries, untreated dental caries, advanced dental caries. Turn the article into a listening activity and create some listening comprehension exercises.
--------------------------------	--

For illustrative purposes, presented in Figures 9-11 are the ChatGPT-generated text and follow-up listening comprehension exercises (multiple choice, T/F, gap-fill and open-ended) that increase ESP students' knowledge of tooth decay and help them acquire collocations focusing on the node *caries*.

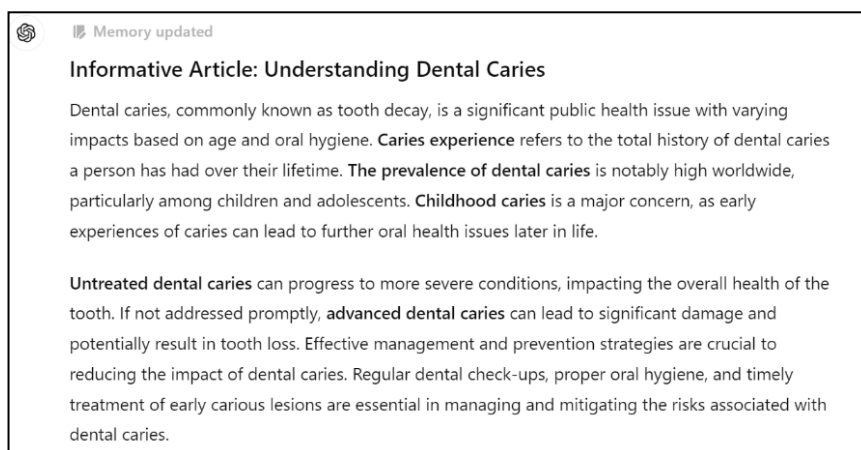


Figure 9: A ChatGPT-generated informative article containing the selected collocates of the node *caries*

Listening Comprehension Exercises
1. Multiple Choice Questions
1. What does "caries experience" refer to?

- A. The current state of dental health
- B. The total history of dental caries a person has had
- C. The frequency of dental check-ups

2. Which group is particularly affected by the high prevalence of dental caries?

- A. Adults
- B. Elderly
- C. Children and adolescents

3. What is a potential consequence of untreated dental caries?

- A. Improved oral health
- B. Advanced dental caries
- C. Reduced risk of tooth loss

Figure 10: A ChatGPT-powered multiple choice listening exercise targeting the selected collocates of the node *caries*

2. True or False
1. Childhood caries do not have long-term effects on oral health. (True / False)
2. Advanced dental caries can lead to tooth loss if not treated. (True / False)
3. Caries experience is an indicator of a person's current level of dental hygiene. (True / False)
3. Fill in the Blanks
1. _____ refers to the history of dental caries a person has had over their lifetime.
2. _____ is a major concern, especially for young children, as it can lead to further dental problems later in life.
3. If dental caries are left _____, they can progress to more severe forms such as advanced dental caries.
4. Regular check-ups and proper oral hygiene are crucial for managing and preventing _____.
4. Short Answer
1. Why is it important to address dental caries promptly?
2. What are two strategies mentioned for managing dental caries effectively?

Figure 11: ChatGPT-powered True/False, gap-fill, and open-ended listening exercises targeting the selected collocates of the node *caries*

Once again, the listening comprehension exercises depicted here are quite similar in content although the text length has been extended to 200 words. Thus, if a wider array of comprehension questions is desired in terms of formulation or content, each comprehension exercise should be brought into focus separately with the aim of regenerating and making tweaks to the AI-generated output.

Besides, it is noteworthy that *caries* represents a technical term, more commonly known as *cavity* or *tooth decay*, to which ESP students' attention should be drawn, especially with regard to rehearsing dentist-patient communication. Still, the technical term *caries* was included here as part of the listening comprehension exercise.

3.3 Selecting appropriate ELT methodologies with the help of ChatGPT

As suggested in *Cambridge GenAI Idea Pack* (2024), GenAI tools lend themselves to determining the most effective language teaching methodologies. Accordingly, ChatGPT can be leveraged when deciding upon the most suitable instructional strategies that could optimise teaching a certain language item. The example below (*Revised prompt 4*) shows how the prompt entitled *Methodology suggestions* was revised based on the original prompting model provided in *Cambridge GenAI Idea Pack* (2024, p. 39) to extend ESP students' knowledge of periodontal health through collocational practice.

Revised prompt 4	Suggest the best ELT methodologies to teach the following collocations to intermediate learners of English for Healthcare at a tertiary level: periodontal disease, periodontal pockets, periodontal status, periodontal therapy, periodontal pocket depth.
-----------------------------	---

When experimenting with a prompt like this, it is crucial to lay emphasis on language learners' proficiency level and their learning context for the sake of obtaining the best-suited output.

Figure 12 and Figure 13 delineate how ChatGPT responded when prompted to propose the best pedagogical approaches to teaching the abovementioned dental collocations to our target student population. Task-based learning (TBL) and communicative language teaching (CLT), more broadly, and case studies, problem-solving tasks, role plays and discussion groups, more specifically, were listed as a practicable means of engaging with our target collocations in the field of periodontal health. It should be noted that ChatGPT provided three additional methodology suggestions presupposing the use of authentic materials that required no involvement on the instructor's part in terms of materials development. However, these suggestions were excluded since they were not in line with the aim of our study concerned with illustrating how GenAI and DDL can be converged for ESP materials development. Instead, the focus was placed on the TBL and CLT approaches, thus allowing the instructors to create the foregoing classroom materials.

2. Task-Based Learning (TBL)

TBL is an effective methodology that focuses on completing meaningful tasks using the target language. This can be applied as follows:

- **Case Studies:** Present students with a case study involving a patient with periodontal issues. Have them discuss the patient's **periodontal status**, the presence of **periodontal pockets**, and the **periodontal therapy** options available.
- **Problem-Solving Tasks:** Assign tasks where students must diagnose a fictional patient's periodontal condition, discuss **periodontal pocket depth**, and recommend appropriate treatments, using the target collocations.

Figure 12: ChatGPT-generated methodology suggestions in line with the TBL approach targeting the selected collocates of the node *periodontal*

4. Communicative Language Teaching (CLT)

CLT emphasizes interaction as both the means and the ultimate goal of learning:

- **Role-Plays:** Create scenarios where students role-play as dentists explaining periodontal issues to a patient. They must use the collocations (e.g., "Your **periodontal status** shows deep **periodontal pockets**...").
- **Discussion Groups:** Organize group discussions on the importance of **periodontal therapy** in preventing **periodontal disease**, encouraging students to use the collocations actively.

Figure 13: ChatGPT-generated methodology suggestions in line with the CLT approach targeting the selected collocates of the node *periodontal*

The ensuing subsections of this paper are comprised of illustrative examples of how methodology suggestions generated by ChatGPT can be actualised in an ESP classroom.

3.4 Creating ChatGPT-powered case studies

Self-created prompt 1 illustrates how ChatGPT was prompted to generate a case study and follow-up discussion questions aimed at instructing a mixed-ability class of ESP university students in diagnosing and managing periodontal disease, with a special focus on the target collocations. The following prompt, as well as those presented in the remainder of the paper, were crafted by the current authors drawing on previous experience of experimenting with this AI-powered chatbot and the prompting models given in *Cambridge GenAI Idea Pack* (2024).

Self-created prompt 1	Generate a case study related to periodontal health for intermediate learners of English for Healthcare at a tertiary level, include the following collocations, highlight them and create additional questions that could help students discuss the case study: periodontal disease, periodontal pockets, periodontal status, periodontal therapy, periodontal pocket depth.
-----------------------	---

Figure 14 and Figure 15 depict a ChatGPT-generated case study dealing with the foregoing subject matter, along with the follow-up discussion questions involving collocations pertaining to the node *periodontal*.

 **Case Study: Periodontal Health**

Patient Profile: Name: John Smith
Age: 45
Occupation: Office Manager
Medical History: Smoker (20 years), occasional drinker
Dental History: Irregular dental check-ups, last visit two years ago

Case Overview: John Smith has recently noticed bleeding gums and bad breath. Concerned, he visits his dentist for a check-up. During the examination, the dentist observes signs of **periodontal disease**. The dentist notes that John has several **periodontal pockets** around his molars, with varying **periodontal pocket depth**. The deepest pockets measure around 5 mm, which indicates a more advanced stage of periodontal issues.

The dentist explains to John that his **periodontal status** is concerning due to the depth of the pockets and the presence of inflammation. The dentist recommends a comprehensive **periodontal therapy** plan, including deep cleaning (scaling and root planing), lifestyle changes like quitting smoking, and more frequent dental visits to monitor the condition.

Figure 14: A ChatGPT-generated case study containing the selected collocates of the node *periodontal*

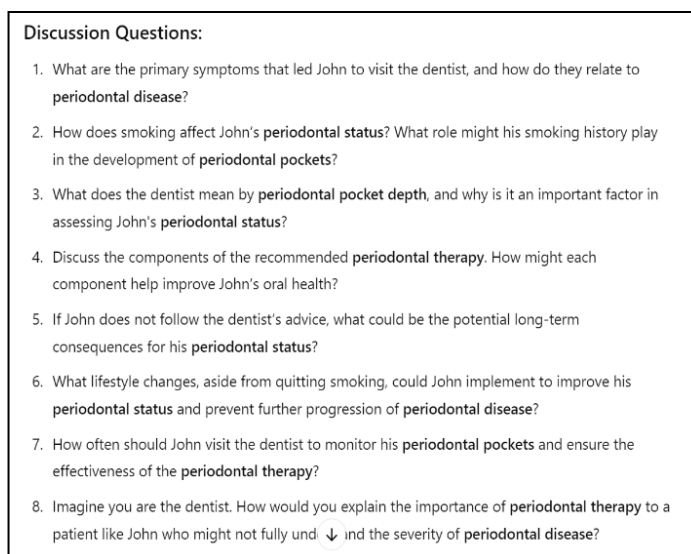


Figure 15: ChatGPT-generated case study discussion questions targeting the selected collocates of the node *periodontal*

As previously mentioned, an expert in the field of dental medicine confirmed the accuracy, relevance, and reliability of the ChatGPT-generated content presented throughout this paper. Provided that the AI-generated output is accurate, relevant and reliable, there is a good chance that a need will not arise to introduce a considerable number of modifications to the textual output itself. What takes a central place in such and similar instances is precisely this convergence of linguistic and content-based knowledge of different subject specialists.

The value of this task lies in its potential to spark discussion and, hence, encourage ESP students to practise their speaking, creative and critical thinking skills. Moreover, ESP students may be asked to draft and act out doctor-patient dialogues based on the AI-generated case study and thus bring the ChatGPT textual output to life.

3.4 Creating ChatGPT-powered problem-based tasks

Self-created prompt 2 demonstrates how the help of ChatGPT was enlisted to create a problem-based task that is intended for a mixed-ability class of ESP university students and directed at broadening their understanding of periodontitis through collocational practice.

Self-created prompt 2	Generate a problem-based task for intermediate learners of English for Healthcare at a tertiary level where students must diagnose a fictional patient's periodontal condition, discuss periodontal pocket depth, and recommend appropriate treatments using the target collocations: periodontal disease, periodontal pockets, periodontal status, periodontal therapy, periodontal pocket depth. Please provide detailed answers to each part of the task, incorporating the target collocations and justifying your recommendations with evidence-based practices.
------------------------------	---

Prompting ChatGPT to respond in detail to each section of the task may prove extremely useful especially for those ESP teachers who grapple with integrating the subject-specific content that is beyond their research scope into classroom activities of this kind. Understandably, the ChatGPT-generated answers need to be approved by an expert in the field, which definitely contributes to considerable strides being made in reinforcing closer cooperation among professionals having expertise in different areas of study.

Figures 16-21 portray a ChatGPT-generated problem-based task covering the topic of periodontal health, accompanied by exemplar responses to each part thereof and targeted at collocations containing the node *periodontal*.


Problem-Based Task

Scenario:

You are a dental healthcare professional evaluating a fictional patient, Emily Turner, who is 38 years old. Emily has come to your clinic with complaints of swollen gums, persistent bad breath, and occasional bleeding while brushing her teeth. During the examination, you find several **periodontal pockets** around her molars. The **periodontal pocket depth** ranges from 4 mm to 6 mm. Her **periodontal status** suggests that she may be suffering from **periodontal disease**.

Your Task:

1. **Diagnosis:**
 - Based on Emily's symptoms and the examination findings, diagnose her periodontal condition. Describe how the **periodontal pocket depth** and the presence of **periodontal pockets** indicate the severity of her **periodontal disease**.

Figure 16: A ChatGPT-generated problem-based task containing the selected collocates of the node *periodontal* (Part 1: Scenario & Diagnosis)

2. Treatment Recommendations:

- Recommend an appropriate **periodontal therapy** plan for Emily. Include both immediate treatments and long-term management strategies to address her **periodontal disease**. Specify how the therapy will help reduce **periodontal pocket depth** and improve her overall **periodontal status**.

3. Follow-Up Plan:

- Design a follow-up plan to monitor Emily's progress. Detail the steps you will take to assess changes in **periodontal pocket depth** and her overall **periodontal status**. Explain how frequently she should return for check-ups and any additional treatments that might be needed based on her response to initial therapy.

4. Patient Communication:

- Draft a brief explanation to present to Emily. Clearly explain her diagnosis of **periodontal disease**, the importance of the **periodontal therapy** you recommend, and how it will benefit her. Ensure the language is accessible and addresses her concerns about treatment and follow-up.

Figure 17: A ChatGPT-generated problem-based task containing the selected collocates of the node *periodontal* (Part 2: Treatment recommendations, Follow-up plan, Patient communication)

 Detailed Answers

1. Diagnosis:

Emily Turner's symptoms of swollen gums, bad breath, and bleeding while brushing, combined with the clinical findings of **periodontal pockets** and a **periodontal pocket depth** ranging from 4 mm to 6 mm, indicate that she is likely suffering from **periodontal disease**.

Periodontal pockets are spaces that form between the teeth and gums due to the destruction of the supporting tissue. The **periodontal pocket depth** is a crucial indicator of the severity of **periodontal disease**. Depths of 4 mm or more typically signify moderate to severe disease, as healthy gums generally have a pocket depth of 1-3 mm. The deeper the pockets, the more advanced the **periodontal disease** is likely to be, affecting the overall **periodontal status**.

Figure 18: ChatGPT-generated exemplar responses to the problem-based task targeting the selected collocates of the node *periodontal* (Part 1: Diagnosis)

2. Treatment Recommendations:

To manage Emily's **periodontal disease**, the following **periodontal therapy** plan is recommended:

- Scaling and Root Planing:** This non-surgical deep cleaning procedure helps remove plaque and tartar from the **periodontal pockets** and smooths the tooth roots to encourage gum reattachment. This can reduce **periodontal pocket depth** and improve **periodontal status**.
- Antibiotic Therapy:** Systemic or local antibiotics may be prescribed to control bacterial infection in the **periodontal pockets**.
- Gum Surgery:** If scaling and root planing are insufficient, surgical options such as flap surgery might be needed to reduce **periodontal pocket depth** and access deeper pockets.
- Oral Hygiene Instructions:** Educate Emily on effective brushing and flossing techniques to prevent further plaque buildup. Emphasize the importance of maintaining good oral hygiene to support the **periodontal therapy** and sustain improvements in **periodontal status**.

Figure 19: ChatGPT-generated exemplar responses to the problem-based task targeting the selected collocates of the node *periodontal* (Part 2: Treatment recommendations)

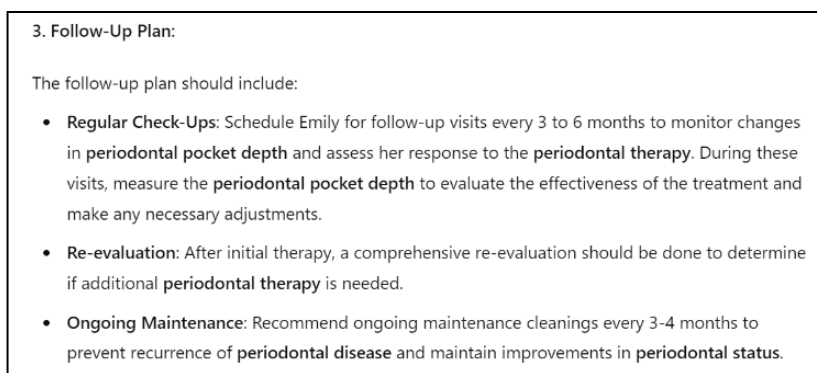


Figure 20: ChatGPT-generated exemplar responses to the problem-based task targeting the selected collocates of the node *periodontal* (Part 3: Follow-up plan)

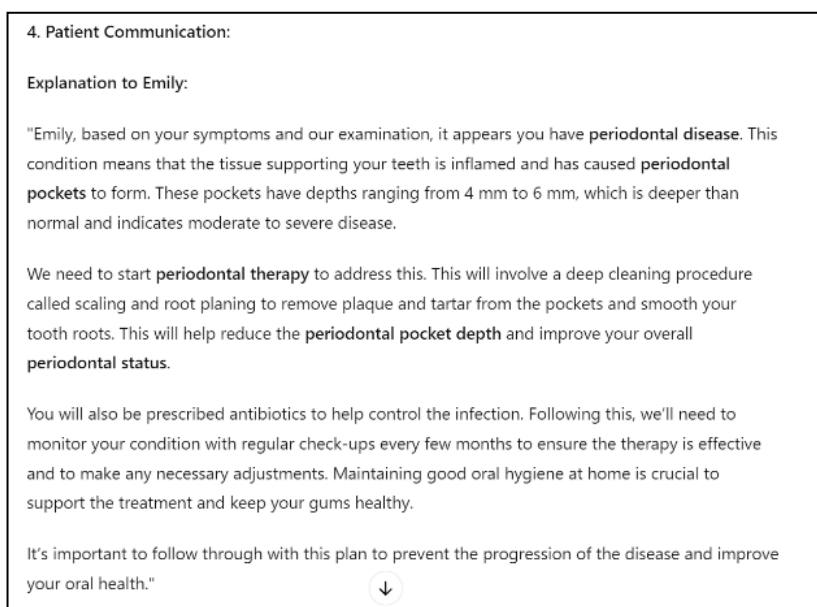


Figure 21: ChatGPT-generated exemplar responses to the problem-based task targeting the selected collocates of the node *periodontal* (Part 4: Patient communication)

This type of task in particular lends itself to honing higher-order thinking skills, apart from the productive ones. Therefore, incorporating a problem-based scenario into ESP instruction may be conducive to deepening ESP students' critical engagement with AI-generated content. To this end, ESP students can be asked, first of all, to solve the task themselves and subsequently compare their results to ChatGPT-generated outputs such as those presented above – which would, in turn, make them less likely to take AI-generated content at face value and, thereby, adopt a more critical stance towards AI tools.

3.4 Creating ChatGPT-powered problem-based role-plays

In line with the aforementioned ChatGPT-generated methodology suggestions, *Self-created prompt 3* shows how ChatGPT was prompted to generate dentist-patient role-plays with a view to helping a mixed-ability class of ESP university students improve speaking skills, while being immersed in the subject-specific content on dental care and related collocational input.

Self-created prompt 3	Generate role-play scenarios for intermediate learners of English for Healthcare at a tertiary level, and include the following collocations: dental caries, dental care, dental services, dental treatment, dental practitioners.
-------------------------------------	---

Figures 22-25 exemplify possible ChatGPT-generated role-play scenarios addressing various aspects of dental care, with a particular emphasis on the target collocations.

Scenario 2: New Patient Consultation Roles: • Patient: Mark • Dental Practitioner: Dr. Patel Situation: Mark has just moved to the area and is visiting Dr. Patel for the first time. He wants to understand the range of dental services offered. Dialogue: • Mark: "Hi Dr. Patel, I'm new to the area and looking for a dentist. Can you tell me about the dental services you offer?" • Dr. Patel: "Welcome, Mark. We provide a comprehensive range of dental services, including preventive care, restorative procedures, and specialized treatments. If you have any specific concerns or need dental treatment, we're here to help." • Mark: "I've had some issues with dental caries in the past. Will your clinic be able to manage that?" • Dr. Patel: "Absolutely. We'll assess your current dental care routine and provide the necessary dental treatment to address any issues."
--

Figure 22: A ChatGPT-generated role-play scenario targeting the selected collocates of the node *dental*
 (Part 1: New patient consultation)

Scenario 3: Treatment Plan Discussion

Roles:

- Patient: Linda
- Dental Practitioner: Dr. Wong

Situation: Linda has been diagnosed with several cavities and needs to discuss her treatment options with Dr. Wong.

Dialogue:

- Linda: "Dr. Wong, I've been told I have several cavities. What are the options for **dental treatment**?"
- Dr. Wong: "We have a few options depending on the extent of the **dental caries**. We'll start with a thorough cleaning and then proceed with fillings or crowns if necessary. It's important to follow a good **dental care** routine to prevent future issues."
- Linda: "How will this treatment affect my overall **dental care**?"
- Dr. Wong: "The treatment will restore your teeth and improve your overall **dental health**. Following our recommended **dental services** will help maintain the results."

Figure 23: A ChatGPT-generated role-play scenario targeting the selected collocates of the node *dental*
(Part 2: Treatment plan discussion)

Scenario 4: Emergency Dental Visit

Roles:

- Patient: James
- Dental Practitioner: Dr. Johnson

Situation: James is experiencing severe tooth pain and needs urgent care. He's visiting Dr. Johnson for emergency **dental treatment**.

Dialogue:

- James: "Dr. Johnson, I'm in a lot of pain and need immediate **dental treatment**. Can you help?"
- Dr. Johnson: "Of course, James. I'll need to examine your teeth to determine the cause of the pain. It might be due to **dental caries** or another issue. We'll provide the necessary **dental care** to relieve your symptoms."
- James: "What should I expect during the treatment?"
- Dr. Johnson: "We'll address the immediate problem and discuss a follow-up plan. Our **dental services** are designed to provide both emergency and ongoing care."

↓

Figure 24: A ChatGPT-generated role-play scenario targeting the selected collocates of the node *dental*
(Part 3: Emergency Dental Visit)

Scenario 5: **Discussing Preventive Measures**

Roles:

- **Patient:** Sarah
- **Dental Practitioner:** Dr. Green

Situation: Sarah is concerned about preventing **dental caries** and wants to learn more about preventive **dental care** strategies.

Dialogue:

- **Sarah:** "Dr. Green, I want to make sure I'm doing everything I can to prevent **dental caries**. What can I do?"
- **Dr. Green:** "Great question, Sarah. Maintaining a good **dental care** routine is crucial. This includes regular brushing and flossing, along with periodic check-ups. We offer various **dental services** designed to prevent and address issues early."
- **Sarah:** "Are there any specific **dental treatments** I should consider for prevention?"
- **Dr. Green:** "We may recommend fluoride treatments or sealants as part of your preventive care plan. This helps protect your teeth from decay and keeps your **dental health** in check."

Figure 25: A ChatGPT-generated role-play scenario targeting the selected collocates of the node *dental*
(Part 4: Discussing preventive measures)

ESP students can benefit not only from acting out role-play scenarios and practising their productive skills, but also from becoming more critically and creatively engaged with AI-generated content – which can be realised, for instance, by producing their own role-play versions and contrasting them with those generated by ChatGPT, comparing several scenarios and choosing the best one while discussing both strong and weak points therein, creating follow-up conversations, etc. Consequently, the key consideration is to involve ESP students in the editing process so that they could assume the role of active participants rather than passive recipients.

4. Concluding remarks

Merging DDL and GenAI not only enables creating more learner-friendly corpus-assisted classroom activities, but it also strengthens collaboration between language teachers and content-based subject specialists in today's AI-driven world. Essentially, it should be underlined that both pre- and post-editing processes are crucial when utilising GenAI tools for ESP materials design purposes, inasmuch as without the human touch, i.e. critical evaluation and refinement, the effectiveness of the AI-generated output is reduced.

A limitation of this research is that no student perceptions regarding the implementation of this research-backed approach to designing ESP materials in the field of dental medicine have been obtained. This could be perceived as one of the significant

implications for future research. What is more, it would be thought-provoking to learn how the convergence of DDL and GenAI presented in this paper could be implemented across various domains of ESP.

References

- Anthony, L. (2024). AntCorGen (Version 1.3.0) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software/AntCorGen>
- Anthony, L. (2025a). AntConc (Version 4.4) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software/AntConc>
- Anthony, L. (2025b). Integrating AI technology into corpus-based language learning through ChatAI. *Computer Assisted Language Learning*, 1–19. <https://doi.org/10.1080/09588221.2025.2589747>
- Asano, M. (2018). Construction of medical research article corpora with AntCorGen: Pedagogical implications. *English Corpus Studies*, 25, 101–115. https://jaecs.com/jnl/ECS25/ECS25_101-115.pdf
- Baisa, V. (2010). Towards disambiguation of word sketches. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, speech and dialogue: TSD 2010 (Lecture Notes in Computer Science*, Vol. 6231). Springer. https://doi.org/10.1007/978-3-642-15760-8_6
- Boulton, A., & Pérez-Paredes, D. (in press). *Data-driven learning: Pedagogy and technology*. In R. Hampel & U. Stickler (Eds.), *The Bloomsbury handbook of language learning and technology*. Bloomsbury Publishing.
- Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics*, 3(3), 100066. <https://doi.org/10.1016/j.acorp.2023.100066>
- Crosthwaite, P., & Boulton, A. (in press). *DDL is dead? Long live DDL! Expanding the boundaries of data-driven learning*. In H. Tyne, M. Bilger, L. Buscail, M. Leray, N. Curry, & C. Pérez-Sabater (Eds.), *Discovering language: Learning and affordance*. Peter Lang.
- Csomay, E., Jablonkai, R. R., & Sun, H. (2025). *Corpora and AI for inductive learning: Theory and practice*. *Applied Corpus Linguistics*, 5(3), 100165. <https://doi.org/10.1016/j.acorp.2025.100165>
- Edmett, A., Ichaporia, N., Crompton, H., & Crichton, R. (2023). *Artificial intelligence and English language teaching: Preparing for the future* (1st ed.). British Council. Retrieved March 24, 2025, from

https://www.researchgate.net/publication/376649238_Artificial_intelligence_and_English_language_teaching_Preparing_for_the_future

- Farooqui, A. (2025). Learning collocations via a data-driven approach: ESP students' perspectives. *Discover Education*, 5, 17. <https://doi.org/10.1007/s44217-025-01048-7>
- Friginal, E., & Prado, M. (2025). *Corpus-informed materials in language teaching*. In L. McCallum & D. Tafazoli (Eds.), *The Palgrave encyclopedia of computer-assisted language learning*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-51447-0_45-1
- Generative AI Idea Pack for English language teachers. (2024). Cambridge University Press & Assessment. Retrieved March 24, 2025, from <https://www.cambridge.org/elt/blog/2024/05/10/generative-ai-for-teachers-free-idea-pack/>
- Johns, T. F. (1991). *Should you be persuaded? Two samples of data-driven learning materials*. In T. F. Johns & P. King (Eds.), *Classroom concordancing* (ELR Journal, No. 4, pp. 1–16).
- Kilgarrieff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. The Sketch Engine: ten years on. *Lexicography*, 1: 7-36, 2014. <http://www.sketchengine.eu>
- Kovačević, D. (2024). ChatGPT and Teaching and Learning of Writing within ESP - a Friend or a Foe?. *2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH)*, East Sarajevo, Bosnia and Herzegovina, 1–6, <https://doi.org/10.1109/INFOTEH60418.2024.10495982>
- Liu, J., & Ma, Q. (2025). Supporting low-proficiency L2 learners' vocabulary learning with custom GPT-scaffolded corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*, 1–37. <https://doi.org/10.1080/09588221.2025.2539152>
- Mizumoto, A. (2023). Data-driven learning meets generative AI: Introducing the framework of metacognitive resource use. *Applied Corpus Linguistics*, 3(3), 100074. <https://doi.org/10.1016/j.acorp.2023.100074>
- Pérez-Paredes, P., & Boulton, A. (2025). *Data-driven Learning in and out of the Language Classroom*. Cambridge University Press.
- Sketch Engine. (n.d.). User Guide. Lexical Computing. Retrieved July 16, 2026, from <https://www.sketchengine.eu/guide/>

- Sun, A. X., & Mizumoto, A. (2025). *Exploring the barriers to data-driven learning in the classroom: A systematic qualitative synthesis*. *Applied Corpus Linguistics*, 5(2), 100126. <https://doi.org/10.1016/j.acorp.2025.100126>
- UNESCO. (2023). *Guidance for generative AI in education and research*. <https://doi.org/10.54675/EWZM9535>. Retrieved March 24, 2025, from <https://unesdoc.unesco.org/ark:/48223/pf0000386693?locale=en>
- Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, 1–7. <https://doi.org/10.1016/j.jslw.2023.101071>
- Xu, J. (2022). *A historical overview of using corpora in English language teaching*. In R. R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 11–25). Routledge.

Sažetak

INTEGRACIJA GENERATIVNE VEŠTAČKE INTELIGENCIJE I UČENJA ZASNOVANOG NA PODACIMA U SVRHU RAZVOJA NASTAVNIH MATERIJALA U ENGLJESKOM JEZIKU STRUKE

Učenje zasnovano na podacima (engl. *Data-Driven Learning (DDL)*) predstavlja efikasan metodološki pristup u nastavi i učenju jezika koji se zasniva na upotrebi jezičkih podataka dobijenih putem korpusa i korpusnih alata, u kombinaciji sa odgovarajućim didaktičko-metodičkim pristupima (Crosthwaite & Boulton, u štampi). Međutim, sa savremenim razvojem generativne veštačke inteligencije (engl. *Generative Artificial Intelligence (GenAI)*), javila se potreba za integriranjem ova dva pristupa kako bi se premostili postojeći jazovi i prevazišla ograničenja DDL-a (Crosthwaite & Baisa, 2023; Mizumoto, 2023). Stoga se u ovom radu oslanjamo na korpus sastavljen od apstrakata naučnih radova iz oblasti stomatologije (veličina korpusa: 109.269 reči) koji je generisan posredstvom alata AntCorGen. Za analizu korpusa korišćen je softver Sketch Engine, krajnji cilj bio je da se prikaže na koji način se učenje zasnovano na podacima i generativna veštačka inteligencija mogu objediniti u procesu izrade nastavnih materijala koji se mogu koristiti za potrebe engleskog jezika struke. U svrhu osmišljavanja niza vežbi i zadataka namenjenih studentima stomatologije sa različitim nivoima jezičke kompetencije engleskog jezika, kao inspiracija poslužila je publikacija *Cambridge GenAI Idea Pack* koja nudi modele promptova koje nastavnici vrlo lako mogu primeniti u različitim nastavnim kontekstima. Skup nastavnih aktivnosti koji ima za cilj usvajanje kolokacija specifičnih za datu oblast, a koji je prilagođen konkretnom obrazovnom kontekstu naših studenata stomatologije, obuhvata raznovrsne vežbe vokabulara i slušanja (popunjavanje praznina, povezivanje, pitanja sa višestrukim izborom, tačno/netačno), kao i aktivnosti koje se zasnivaju na induktivnom učenju poput studija slučaja, problemski orijentisanih zadataka i igre uloga (engl. *role-play*) stomatologa i pacijenta. Ovakav pristup izradi nastavnih materijala namenjenih studentima engleskog jezika struke donosi praktične implikacije u prilog integriranju korpusnih podataka i generativne veštačke inteligencije kako bi studenti na tercijarnom nivou obrazovanja u što većoj meri bili izloženi autentičnoj upotrebi jezika.

Ključne reči: engleski jezik struke u zdravstvu, igra uloga, kolokacije, korpusna lingvistika, problemski orijentisano učenje, učenje zasnovano na podacima, veštačka inteligencija

Primljeno: 14. 1. 2026.

Prihvaćeno: 30. 9. 2026.