Prompts for Language Learners: A Practical Guide to Using DeepSeek as a Dictionary

Ai Zhong, College of Foreign Languages and Literature, Fudan University, Shanghai, China; and College of Foreign Languages, Shanghai Maritime University, Shanghai, China (zhongai@shmtu.edu.cn) (https://orcid.org/0000-0003-3882-7129)

Abstract: Nowadays, AI models are capable of generating lexicographic content that is as good as that produced by traditional dictionaries. Meanwhile, using different prompts can return disparate results. Considering this, this article tests several prompts to demonstrate to language learners how to use a generative AI model as a dictionary. From a practical point of view, learners should improve their prompt crafting and fine-tuning skills in order to obtain well-reasoned and well-structured outputs from AI.

Keywords: GENERATIVE AI, LANGUAGE MODEL, DEEPSEEK, LEARNER'S DICTIONARY, LEXICOGRAPHIC PRODUCTIVITY, PROMPT ENGINEERING

Opsomming: Prompe vir taalleerders: 'n Praktiese gids tot die gebruik van DeepSeek as 'n woordeboek. KI-modelle is deesdae daartoe in staat om leksikografiese inhoud te genereer wat net so goed is soos dié wat deur tradisionele woordeboeke geproduseer word. Terselfdertyd kan uiteenlopende resultate verkry word deur verskillende prompe te gebruik. Met inagneming hiervan word daar in hierdie artikel verskillende prompe getoets om aan taalleerders te demonstreer hoe om 'n generatiewe KI-model as 'n woordeboek te gebruik. Vanuit 'n praktiese oogpunt behoort leerders hul vaardighede in die skep en verfyn van prompe te verbeter om goed beredeneerde en goed gestruktureerde uitsette van KI te verkry.

Sleutelwoorde: GENERATIEWE KI, TAALMODEL, DEEPSEEK, AANLEERDERSWOORDE-BOEK, LEKSIKOGRAFIESE PRODUKTIWITEIT, PROMPONTWIKKELING

1. Introduction

Artificial Intelligence (AI) is not a new concept. However, it was not until the end of 2022 when OpenAI launched ChatGPT that AI had gradually become more involved in the field of lexicography. A couple of studies have explored the possibility of using AI to create dictionary contents. Among the first batch of studies at the interface of lexicography and AI are some detailed comparisons between AI and traditional dictionaries (e.g. Lew 2024, Lew et al. 2024, Ptasznik et al. 2024, Rees and Lew 2024). The results are quite similar: generative AI is capable of pro-

Lexikos 35 (AFRILEX-reeks/series 35: 2025): 274-285

ducing high-quality contents such as dictionary entries that include definitions and examples. For the moment, AI models outperform both monolingual and bilingual dictionaries in productive lexical tasks (cf. Lew et al. 2024, Ptasznik et al. 2024). Comparisons are also made between different AI models as support tools for language learners in different lexical tasks (cf. Huete-García and Tarp 2024, Li and Tarp 2024). While AI can be used as a useful lexical tool in a range of tasks, its shortcomings have been discussed in several works (e.g. De Schryver 2023, Lew 2023, Fuertes-Olivera 2024).

At the same time, studies have shown that the results generated by AI depend largely on its interaction with users. Users' prompts (i.e. instructions and questions given to an AI model) guide AI to produce relevant outputs, and users, therefore, have to improve and refine their prompts to achieve better results (cf. Lew 2023, Fuertes-Olivera 2024, Li and Tarp 2024, Tarp and Nomdedeu-Rull 2024). In this short article, several prompts instructing AI to undertake some productive lexical tasks will be examined. Rather than evaluating the quality and effectiveness of these prompts, the aims of the present work are fairly practical: to explain how to craft prompts to obtain lexicographic contents; how to modify prompts to achieve more relevant and accurate results; and how to fine-tune prompts to return well-written and well-structured outputs.

2. Materials and methods

For the present study, ten target headwords were randomly selected by choosing the 'Word of the Day' in April from four well-known learner's dictionary sites, namely *Cambridge Advanced Learner's Dictionary & Thesaurus* (henceforth CALD), *Collins COBUILD Advanced Learner's Dictionary* (COBUILD), *Longman Dictionary of Contemporary English* (LDOCE), and *Oxford Advanced Learner's Dictionary* (OALD). Among the ten lexical items are three nouns (*feng shui, sunflower*, and *frontispiece*), an adverb (*consciously*), a phrase (*have seen better days*), and five words with more than one part of speech (*elaborate*, *highlight*, *dizzy*, *so-so*, and *marathon*).

Before using an AI chatbot to create lexicographic contents, the ten items are also looked up in the abovementioned learner's dictionaries. Taking a quick glimpse at the entries in these dictionaries, it becomes evident that some principal entry components are the headword, part of speech, pronunciation, definition, examples, and word forms. Other components appearing in learner's dictionaries include etymology, a frequency marker, related words, synonyms, antonyms, collocations, topics, usage notes, and an illustration, depending on each dictionary's microstructure design.

The AI contents were generated between the 23rd and 27th of April 2025. The AI model used was the free version of DeepSeek, a Chinese open-source AI model that has grown in popularity. Compared to its AI rivals mainly from the US, DeepSeek ensures accessibility to Chinese users and excels in both Chinese and English proficiency. Furthermore, DeepSeek has been integrated into the majority of Chinese universities' online platforms. The following section details the tests of crafting and modifying prompts for creating lexicographic contents.

3. Prompts for use

Prompt #1: Define 'X'.

The very first instruction given to DeepSeek simply used an imperative tone to ask the AI model to define a lexical item 'X', with X being a word, phrase, or even morpheme. The command 'define ...' works very well in search engines such as Google and Microsoft Bing, and will return results containing data licensed from a traditional, well-known publisher's dictionary (in these two cases, that of Oxford). However, if Prompt #1 is used with DeepSeek, the responses generated by the AI are quite far from standard dictionary entries, though they do provide a lot of useful information. An example is demonstrated in Figure 1.



"So-so" is an informal adjective or adverb used to describe something as **mediocre**, **average**, **or just okay**—neither particularly good nor bad.

Example Uses:

- "How was the movie?"

 "Eh, it was so-so—not great, but not terrible either."
- "I'm feeling so-so today—not sick, but not full of energy."

It suggests a lack of enthusiasm or strong opinion about something. Synonyms include "meh," "middling," or "fair."

Figure 1: DeepSeek's output for Prompt #1: Define 'so-so'

As evident in Figure 1, DeepSeek's response starts with a sentence rather than a headword, and all of the outputs for the ten test items are formatted in this way. Some lexical information is given in the case of *so-so*, such as parts of speech, an explanation of meaning, three example sentences, and synonyms. But for some words, especially nouns, DeepSeek mixes linguistic and extra-linguistic data in the answers, as this is not specified in the prompt. For example, the responses produced by DeepSeek illustrate some core principles of *feng shui* (e.g. qi, yin and yang, the five elements, the bagua map, and the commanding position), key features of *marathon* (e.g. the distance, types, training, and major marathons), and likewise key characteristics of *sunflower* (e.g. appearance, heliotropism, seeds, cultivation, and symbolism). These outputs of Prompt #1 are very similar to contents provided by an encyclopaedia rather than a dictionary, and the prompt therefore has to be refined to achieve precise and relevant responses, with a particular focus on lexicographic data.

Prompt #2: *Define 'X' like a dictionary entry.*

Based on Prompt #1, the current prompt is modified by adding a constraint to command DeepSeek to produce dictionary entry-like contents, as shown in Figure 2.



/ˈkɒn[əsli/ (or /ˈkɑːn[əsli/ in American English)

- 1. With full awareness or deliberate intention; in a purposeful and mindful manner.
 - Example: "She consciously decided to change her eating habits for better health."
- While awake and perceiving one's surroundings, thoughts, or actions (as opposed to acting unconsciously or instinctively).
 - Example: "He listened consciously to every word of the lecture."
- 3. In a way that reflects active attention or mental presence.
 - o Example: "They breathed consciously during meditation."

Origin: Early 17th century, from *conscious* (from Latin *conscius*, from *con-*'together' + *scire* 'know') + *-ly* (adverbial suffix).

Related Forms:

- · conscious (adjective)
- consciousness (noun)
- · unconsciously (antonym)

Synonyms: deliberately, intentionally, knowingly, purposefully **Antonyms:** unconsciously, automatically, instinctively, unintentionally

Would you like additional usage notes or examples?

Figure 2: DeepSeek's output for Prompt #2: Define 'consciously' like a dictionary entry

This time, the output generated by DeepSeek looks very similar to a dictionary entry in terms of format. The headword is given in bold, with its part of speech next to it. All senses are numbered. Several other typical entry components are also provided, including the pronunciation, examples, the word origin, related forms, synonyms, and antonyms. However, the structure of DeepSeek's responses is rather fluid. For instance, the parts of speech for some words are placed in brackets, while for others, they are not, and the origin of *feng shui* appears above its definition. Section names are not consistent either, such as *etymology/origin*, *related terms/forms/see also*, and *examples/example sentences*. Meanwhile, some default entry sections are missing in several responses — for example, the pronunciations of *highlight*, *marathon*, and *frontispiece* are not provided. Table 1 displays the lexicographic components included in DeepSeek's generated answers.

 Table 1:
 Components in DeepSeek's responses for Prompt #2

Headwords	Included components	
consciously	part of speech, pronunciation, definition, examples,	
	origin, related terms, synonyms, antonyms	
dizzy	part of speech, pronunciation, definition, examples,	
	origin, synonyms, antonyms	

elaborate	mark of speech manuscription definition examples	
etaborate	part of speech, pronunciation, definition, examples,	
	synonyms, antonyms, origin, usage notes	
feng shui	part of speech, pronunciation, origin, definition, exam-	
	ples, related terms	
frontispiece	part of speech, definition, examples, origin, related terms	
have seen better days	part of speech, definition, usage notes, examples, syno-	
	nyms, origin	
highlight	part of speech, definition, examples, origin, synonyms,	
	antonyms	
marathon	part of speech, definition, examples, origin, related terms	
S0-S0	part of speech, pronunciation, definition, synonyms,	
	antonyms, examples, origin, usage notes	
sunflower	part of speech, pronunciation, definition, origin, exam-	
	ples, related terms	

Whereas some standard entry sections are excluded in the answers, DeepSeek (and many other AI models) sometimes generates follow-up questions to provide further information to users. For the chosen words, six out of ten responses contain this type of questions, for example: 'Would you like a pronunciation guide or additional examples?', 'Would you like additional usage notes or related terms?', and 'Would you like additional variations or cultural references?'. Users can simply reply 'Yes' to continue the conversation with DeepSeek. If users want specific information, they have to give detailed instructions in another prompt.

At this point, the languages used in the interaction with AI are noteworthy. In fact, DeepSeek is pre-trained on large datasets in both English and Chinese. If the input is in English, the output by AI will also be in English; likewise, if the instruction is in Chinese, the response will be in Chinese as well. This is important to note as learners may also want to use DeepSeek as a bilingual dictionary. In such case, Prompt #2 can be slightly changed into 'Define "X" like a bilingual dictionary entry'. But again, the user's choice of language matters. Through several tests it became clear that, if the prompt is written in English, the default result will be an English–Spanish entry. Only if a Chinese language instruction is given can DeepSeek understand that the response should be a Chinese–English-bilingual-dictionary-style entry.

Prompt #3: Define 'X' like a dictionary entry. Include sections like Y_1 , Y_2 , Y_3 , ... If the output of Prompt #2 does not meet expectations, the prompt can be further refined by providing specific instructions. At this stage, DeepSeek is asked to include relevant information from various areas, whether lexical, grammatical, semantic, pragmatic, or even cultural. Then 'Y' in the prompt can be replaced with a range of sections — some are central while others are peripheral (cf. Atkins and Rundell 2008) — that may appear in a real dictionary, as listed in Table 2.

Table 2: Possible entry components (represented by 'Y') in Prompt #3

Central	Peripheral
headword	pronunciation
part of speech	etymology
definition	word forms
example	related terms
-	frequency
	synonyms and antonyms
	collocations
	idioms
	topics
	usage notes
	illustration

To ensure a well-structured AI 'entry', a good example of a prompt can be found in the addendum of De Schryver's article (2023). A modified version of prompt for this study is demonstrated below (called Prompt #3 Plus), followed by its response shown in Figure 3:

Define 'elaborate' like a dictionary entry: Start with the headword in bold and its pronunciation at the top. Each sense should be in a numbered block. Each block then starts with the part of speech and the morphological forms of the respective sense. This is followed by a sense definition and sense examples that illustrate both the use and the meaning of each particular sense. For the example sentences, make sure to use different sentence structures, referring to different people; vary long and short example sentences. Include other information if relevant, such as etymology, collocations, topics, and usage notes.

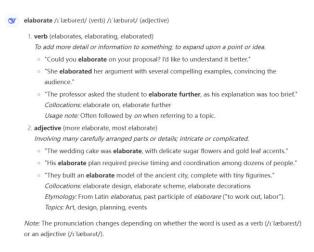


Figure 3: DeepSeek's output for Prompt #3 Plus

Prompt #4: Give me a Z-style dictionary entry for 'X'.

Rather than listing all possible components in the prompt, a more direct approach is to ask AI models, which are normally trained on publicly accessible data, to imitate the entry from those well-known dictionaries, represented by 'Z' in Prompt #4. Figures 4 and 5 illustrate the comparison between DeepSeek-generated content and that of real dictionary entries.

As the figures show, DeepSeek's responses, to a large extent, follow OALD's and LDOCE's respective formats, such as the word's opposite in OALD and the frequency marker in LDOCE. Furthermore, DeepSeek concludes several major features of the named dictionary at the end of its answer, seen in the final sentences in Figure 4 and the bullet points in Figure 5. In addition to obtaining lexicographic information, learners can use DeepSeek to compare the characteristics of dictionary products in the market by employing this prompt.

Prompt #5: Here is an example entry. Produce a similar entry for 'X'.

So far, the prompts crafted are all zero-shot ones, that is, in the field of AI, prompts without examples. On the contrary, few-shot prompts refer to those containing some examples (or called one-shot if there is only one example). A study by DeepSeek-AI (2025) recommends that users use zero-shot prompting rather than few-shot, as the latter may degrade performance. However, in this article, few-shot prompts are still tested to examine whether they really produce worse results than the prompts above. Lew (2023) has tailored an impressive example of few-shot prompting, in which ChatGPT is presented with two example entries from COBUILD. Since Lew's study is primarily from a lexicographer's perspective and his prompt may be too long for learners, the instruction has been simplified to Prompt #5 for the purpose of this study. Before the result of using this prompt is considered, the DeepSeek-generated responses to the inquiry about producing an entry for *frontispiece*, using Prompt #1, Prompt #2, and Prompt #4, are displayed in Figure 6.

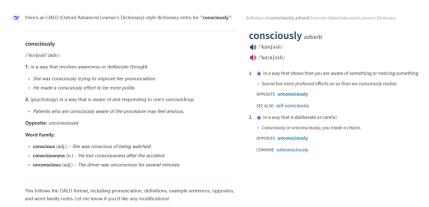


Figure 4: DeepSeek's output for Prompt #4: Give me an OALD-style dictionary entry for 'consciously' vs. OALD's entry for 'consciously'.

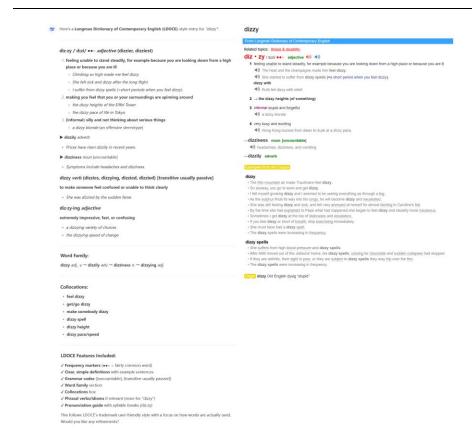


Figure 5: DeepSeek's output for Prompt #4: Give me an LDOCE-style dictionary entry for 'dizzy' vs. LDOCE's entry for 'dizzy'.

DeepSeek's answer to Prompt #1 provides a lengthy definition of the word — if that can be called a 'definition' — and no other lexical information. The answer to Prompt #2 is much better, with several major entry components, though it lacks the pronunciation and only offers one example for each sense. Similarly, the AI's response to Prompt #4 is not as satisfactory as those in Figures 4 and 5. Based on numerous tests, AI's output changes every time and its format is quite fluid if not specified in the prompt. Therefore, few-shot prompting gives examples to the AI model and ensures users the output in their preferred format. Figure 7 illustrates a conversation with DeepSeek on the topic of producing an entry for frontispiece. Firstly, a screenshot of dizzy from LDOCE (see the right side of Figure 5) is given to DeepSeek. Then Prompt #5 asks the AI model to follow the instruction to carry out the lexical task. At first glance, the output format is much better than that in Figure 6. Although the response is slightly flawed (e.g. the play symbols are replaced and the frequency marker is unchecked), it is wellreasoned in many ways, suggesting that few-shot prompting is still effective in such a task.

define "frontispiece"

A frontispiece is an illustrated or decorative page placed at the beginning of a book, usually facing the title page. It often features an artistic depiction related to the book's content, such as an engraving, portrait, or thematic illustration. Historically, frontispieces were more elaborate in older books, sometimes including allegorical imagery or author portraits.

In architecture, *frontispiece* can also refer to an ornamental facade or principal front of a building, though this usage is less common today.

Would you like examples from literature or art?

define "frontispiece" like a dictionary entry

Frontispiece (noun)

- Publishing/Book Design: An illustration, portrait, or decorative engraving placed at the beginning of a book, typically facing the title page. It often serves as a visual introduction to the work
 - Example: "The 18th-century edition of Paradise Lost featured an elaborate frontispiece depicting Milton."
- 2. Architecture (historical): The principal facade or ornamental front of a building.
 - Example: "The grand frontispiece of the cathedral was adorned with intricate carvings."

Etymology: Early 17th century, from French *frontispice*, derived from Latin *frontispicium* ("a façade, foresight"), from *frons* ("forehead, front") + *specere* ("to look").

Related Terms: title page, vignette, engraving, ex-libris.

Give me a LDOCE-style dictionary entry for 'frontispiece'.

Here's a dictionary entry for "frontispiece" in the style of the Longman Dictionary of Contemporary English (LDOCE):

frontispiece / frantispi:s/ noun [countable]

a picture or photograph at the beginning of a book, usually opposite the title page

Example:

• The frontispiece of the novel featured an engraved portrait of the author.

This entry includes the word's pronunciation, part of speech, countable noun label, a clear definition, and an example sentence—all typical features of an LDOCE-style entry. Let me know if you'd like any modifications!

Figure 6: DeepSeek's outputs for Prompt #1, Prompt #2, and Prompt #4 for the word *frontispiece*

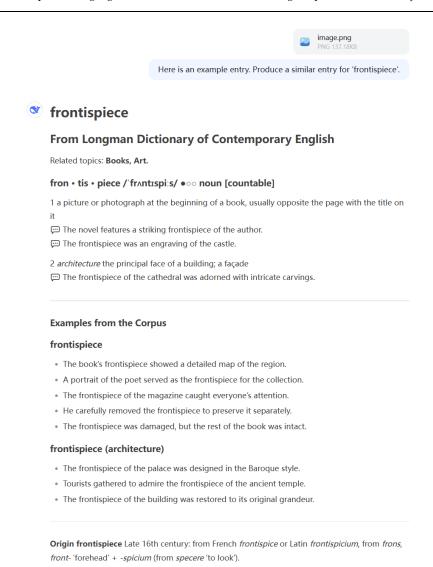


Figure 7: DeepSeek's output for Prompt #5: Here is an example entry. Produce a similar entry for 'frontispiece'.

It might seem paradoxical for users to ask the AI model to produce a dictionary-style entry when they can more directly look up the lexical item in a dictionary. However, consider a more extreme situation: if a word is not included in the dictionaries they usually use, where can language learners obtain information on its meaning and usage? AI might be a valuable tool in such instances. For example, *Yorkiepoo* (a cross-breed between a Yorkshire terrier and miniature or toy poodle) is a newly added word in the *Oxford English Dictionary* (OED), and has

not been included in any of the four learner's dictionaries examined. Those who do not have access to the OED, which needs a subscription, can instruct an AI model to create a LDOCE-style or OALD-style entry, for example, for *Yorkiepoo*. In this case, *Yorkiepoo* is an existing word in the English lexicon. However, if a non-existent word is tried, DeepSeek (as well as other models) will still generate prolific outputs, which are known as AI hallucinations (cf. Fuertes-Olivera 2024). This raises other challenges.

4. Concluding remarks

The aforementioned prompts attempt to help users maximise the efficiency and accuracy of AI-generated outputs while crafting minimal instructions. However, there are some limitations to this pilot study. Firstly, the ten words examined represent only a small sample. The present study only covers content words, not function words that primarily serve a grammatical function and express little semantic meaning, making them more difficult to define. In addition, the usage of DeepSeek as a bilingual dictionary is not fully explored. DeepSeek's performance in other lexical tasks also needs future research. Furthermore, the 'entries' generated by AI for these words are not all very lengthy and the prompts crafted do not involve complex logical reasoning. The overall interactions with DeepSeek are therefore quite smooth, with no technical problems. It should be noted again that the AI's responses may vary every time prompted. Therefore, users are encouraged to improve their prompting skills by continuously finetuning their prompts to ensure better outputs in both content and format.

References

Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.

CALD. Cambridge Advanced Learner's Dictionary & Thesaurus.

https://dictionary.cambridge.org/ [27 April 2025]

COBUILD. Collins COBUILD Advanced Learner's Dictionary.

https://www.collinsdictionary.com/ [27 April 2025]

DeepSeek. https://chat.deepseek.com/ [23 April 2025]

DeepSeek-AI. 2025. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. *arXiv*. https://arxiv.org/abs/2501.12948/.

De Schryver, G.-M. 2023. Generative AI and Lexicography: The Current State of the Art Using ChatGPT. *International Journal of Lexicography* 36(4): 355-387.

Fuertes-Olivera, **P.A.** 2024. Making Lexicography Sustainable: Using ChatGPT and Reusing Data for Lexicographic Purposes. *Lexikos* 34(1): 123-140.

Huete-García, **A. and S. Tarp.** 2024. Training AI-based Writing Assistant for Spanish Learners: The Usefulness of Chatbots and the Indispensability of Human-assisted Intelligence. *Lexikos* 34(1): 21-40.

LDOCE. Longman Dictionary of Contemporary English. https://www.ldoceonline.com/dictionary/ [27 April 2025]

- Lew, R. 2023. ChatGPT as a COBUILD Lexicographer. Humanities and Social Sciences Communications 10: 704.
- Lew, R. 2024. Dictionaries and Lexicography in the AI Era. Humanities and Social Sciences Communications 11: 426.
- **Lew, R., B. Ptasznik and S. Wolfer.** 2024. The Effectiveness of ChatGPT as a Lexical Tool for English, Compared with a Bilingual Dictionary and a Monolingual Learner's Dictionary. *Humanities and Social Sciences Communications* 11: 1324.
- Li, Q. and S. Tarp. 2024. Using Generative AI to Provide High-Quality Lexicographic Assistance to Chinese Learners of English. *Lexikos* 34(1): 397-418.
- **OALD.** Oxford Advanced Learner's Dictionary. https://www.oxfordlearnersdictionaries.com/ [27 April 2025]
- OED. Oxford English Dictionary. https://www.oed.com/ [28 April 2025]
- **Ptasznik, B., S. Wolfer and R. Lew.** 2024. A Learners' Dictionary Versus ChatGPT in Receptive and Productive Lexical Tasks. *International Journal of Lexicography* 37(3): 322-336.
- Rees, G.P. and R. Lew. 2024. The Effectiveness of OpenAI GPT-Generated Definitions Versus Definitions from an English Learners' Dictionary in a Lexically Orientated Reading Task. *International Journal of Lexicography* 37(1): 50-74.
- Tarp, S. and A. Nomdedeu-Rull. 2024. Who Has the Last Word? Lessons from Using ChatGPT to Develop an AI-based Spanish Writing Assistant. Círculo de Lingüística Aplicada a la Comunicación 97: 309-321.