

Using CorpusMate

The diagram shows the CorpusMate interface with several callouts:

- Enter or paste text in this box**: Points to the search input field.
- You can select from subject area/mode filter dropdown**: Points to the 'in topic' dropdown menu.
- Example queries using the operators are provided**: Points to a box containing wildcard and Boolean operators.

Wildcard

Symbol	Meaning
.* (dot asterisk)	next to/within part of a word to find all forms of that word
? (question mark)	to see if a word is necessary or not
/ (slash)	to compare alternative words
"" (quotation marks)	to search for exact word form match
! (exclamation mark)	to search for anything except the word

Example

lik.* to * extent
sit down?
emphasise/emphasize
"book"
jurassic !period

Feedback and bug reports are welcome!

Steps:

1. Enter a Query

Type your search word or phrase into the search bar. You may also select a subject area or mode filter from the dropdown menus to narrow down your results.

The screenshot shows the CorpusMate interface with the search bar containing the text "Nice". The "in topic" dropdown menu is open, showing a list of subject areas. The "Computer Games" option is highlighted.

Subject Area Dropdown Menu:

- Architecture, Planning and Design
- Biology
- Business and Economics
- Chemistry
- Computer Games**
- Culture, Arts and Music
- Education
- Engineering
- English Language and Literature
- Geography, Agriculture and Environment
- Health and Medicine
- History
- Journalism
- Law
- Mathematics
- Other subjects

* **Choosing the topic "Computer Games" is recommended** for exploring language that closely resembles comments found on Reddit. However, students are free to select other topics based on their interests.

2. Use Advanced Search Operators (Optional)

Enhance your search using the following wildcard and Boolean operators:

- Dot asterisk (.*) matches any number of characters (e.g., play.* finds *plays*, *playing*, *played*).

- Question mark (?) shows whether the word is necessary (e.g., color/?colour matches both spellings).
- Slash (/) acts as an "OR" operator to compare alternative words (e.g., happy/sad).
- Quotation marks ("") searches for exact matches (e.g., "online game" will only return that exact phrase).
- Exclamation mark (!) excludes specific terms (e.g., !boring finds all results except those including *boring*).

3. Click 'Search'

Once your query is ready, click the **Search** button to view results from the corpus database.

Exploring Results

The tool provides four main features for learning how words and phrases are used in real contexts:

- Concordance: View the word or phrase in various sentence contexts.
- Compare Results: See differences in usage across different corpora or categories.
- Pattern Finder: Identify common word patterns.
- Distribution: Analyze how frequently a word appears across different topics.

Note: For most learning purposes, especially for identifying how a word is used in context, using the **Concordance** function alone is sufficient.

1. KWIC View / Sentence View

- **KWIC (Key Word in Context):** Displays results in a compact format with the target word (query) in the center. Words around it are shown in colored patterns to highlight common structures.
- Concordance lines are sorted based on frequent 4-grams starting with the query word.
- Only one sample of each 4-gram is shown to avoid repetition; you can click the arrow icon to expand and see all related results.
- **Sentence View:** Shows full sentences containing your query.
- Click 'Expand' to add one sentence before and after the result, up to 10 additional sentences.
- Ideal for deeper context and teaching materials.

**You can switch between views anytime.*

2. All / Spoken / Written (Mode Filter)

Use this option to filter your results by language mode:

- All: Includes both spoken and written language.
- Spoken: Shows only results from conversations, interviews, or speech transcripts.
- Written: Shows results from articles, essays, blogs, etc.

3. All / Academic / General (Topic Filter)

This lets you choose the type of content:

- All: Includes both academic and general content.
- Academic: Focuses on research, scholarly writing, or educational texts.
- General: Includes everyday language, such as blogs and informal writing.

4. Left Patterns / Right Patterns

View common 4-word patterns (4-grams):

- Right Patterns: See what typically follows your query word.
- Left Patterns: See what commonly comes before your query word.

**Click 'Switch to Left/Right Patterns' to toggle views.*

5. Only in Topic (Topic-Specific Search)

- Restrict your search to a specific topic

6. Only in Collection (Sub-Corpus Filter)

- Choose a sub-corpus or collection to limit your search.

7. Show KWIC / Hide KWIC

This toggle lets you:

- Show KWIC: See the target word in the center.
- Hide KWIC: Hide the target word in the center.

Using DeepSeek AI for Concordance Analysis

Why Use DeepSeek AI for Concordance Analysis?

1. No input size limitation: It allows users to paste a large amount of text (such as concordance lines) without restrictions on word or character count.
2. Advanced reasoning: Ideal for analyzing linguistic features such as themes, collocations, and discourse structures.

⚠ Limitations and Privacy Considerations

1. Language limitation: DeepSeek currently supports only English and Chinese for both input and output, which may limit accessibility for users working in other languages.
2. Privacy caution: Users should avoid entering personal, sensitive, or confidential information to ensure data privacy while using the platform.

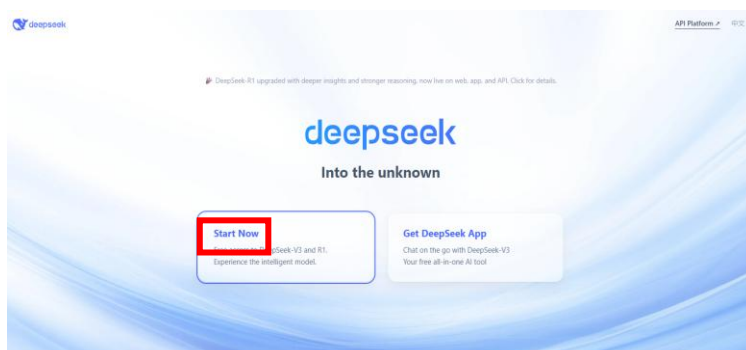
Follow these steps to use DeepSeek AI in analyzing concordance data from CorpusMate:

1. Open your browser and go to: <https://www.deepseek.com/>.
2. Click the English language button at the top-right corner.

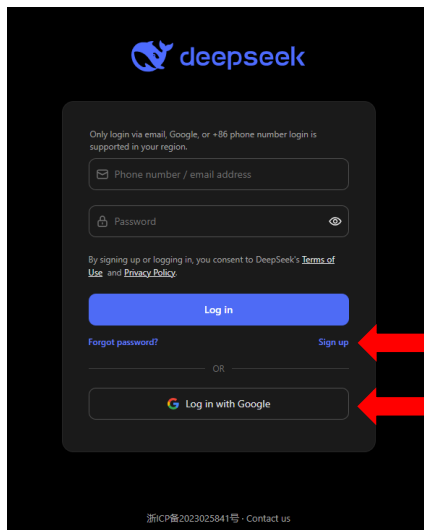


3. Click "Start Now".

- This will direct you to the main platform to begin your session.

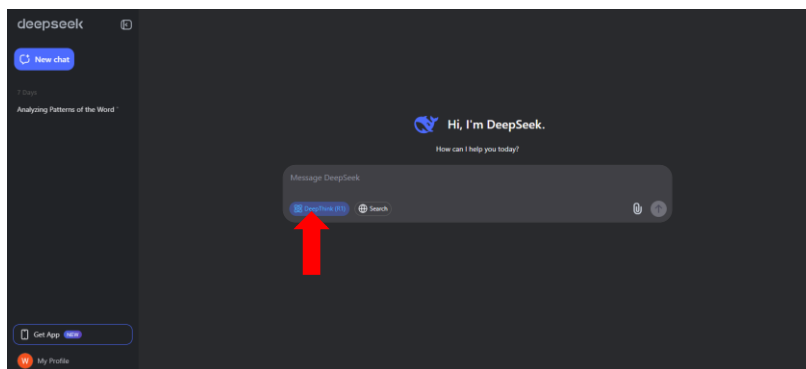


4. Sign up or log in using your Google account (Gmail) for quick access.



5. Click on “DeepThink (R1)”.

- This feature is ideal for tasks involving deep reasoning and analysis.



6. Enter a structured prompt

- You may use the following format for your prompt.

The prompt (input) should be clear, specific, and applicable to analysis tasks. It should include the following components:

Role: Define the persona or role of AI (e.g., a linguistics data analysis expert).

Purpose or Goal: Explain the objective or reason for needing the output.

Context: Describe the relevant situation or background (e.g., for research reporting).

Input: Give details about what you will provide and what the AI should focus on.

Constraints and Guidelines: Mention any limitations or directions (e.g., use only the given concordance lines, avoid assumptions beyond the data)".

Desired Output: State clearly what kind of analysis or result you want.

Prompt Sample 1: Common patterns identification

Role: You are a linguistics data analysis expert specializing in corpus linguistics and discourse analysis.

Purpose or Goal: The analysis will be used as a guideline to support language learning for me as a learner.

Input: I will provide you with XXX concordance lines for the word X. You do not need to analyze all XXX lines. Instead, focus only on the most important lines that clearly show how the word is commonly used.

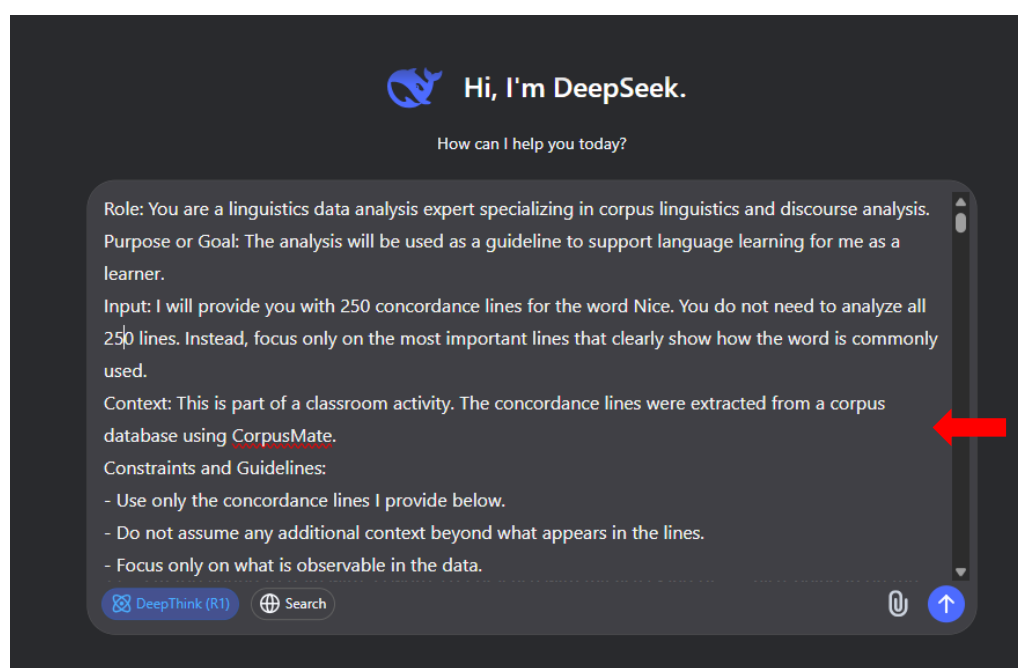
Context: This is part of a classroom activity. The concordance lines were extracted from a corpus database using CorpusMate.

Constraints and Guidelines:

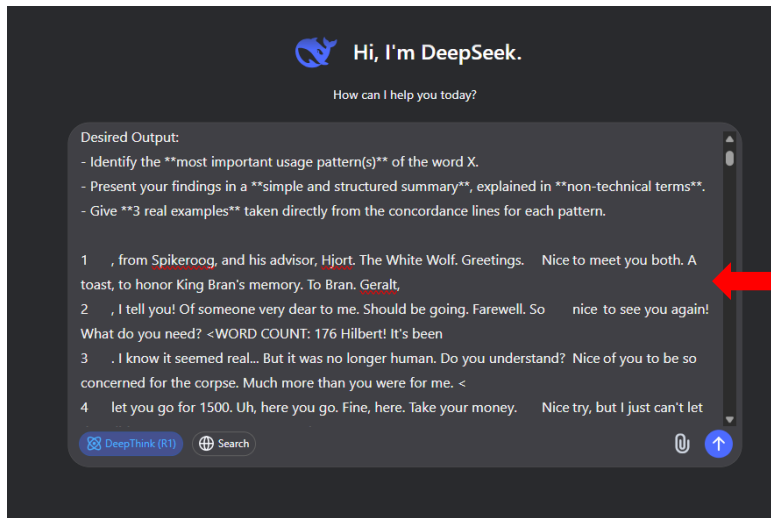
- Use only the concordance lines I provide below.
- Do not assume any additional context beyond what appears in the lines.
- Focus only on what is observable in the data.
- Identify frequent and salient patterns of usage.
- Avoid making up examples or interpreting beyond the data given.

Desired Output:

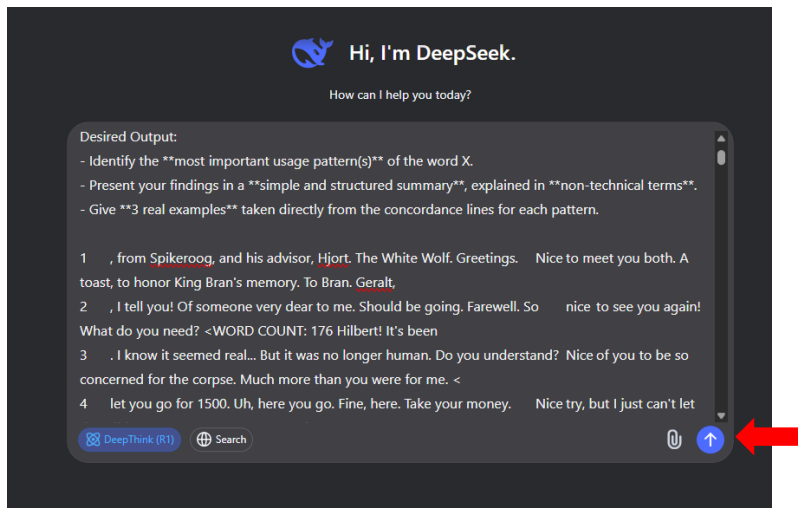
- Identify the ****most important usage pattern(s)**** of the word X.
- Present your findings in a ****simple and structured summary****, explained in ****non-technical terms****.
- Give ****3 real examples**** taken directly from the concordance lines for each pattern.



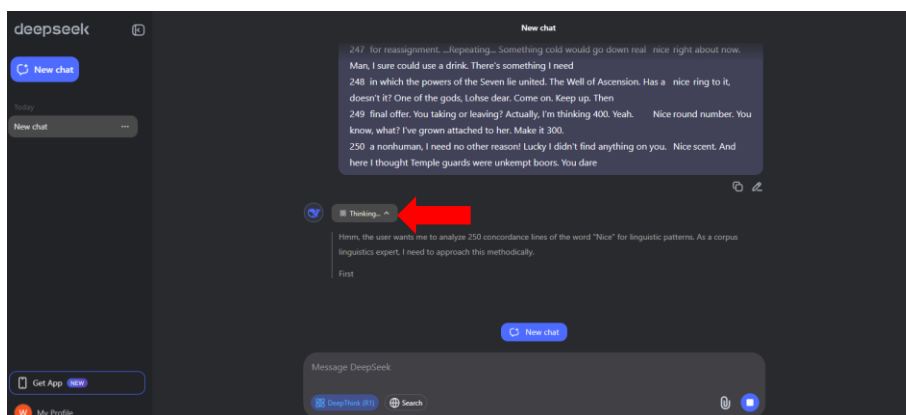
7. Copy concordance lines from CorpusMate and paste them *after* the structured prompt.



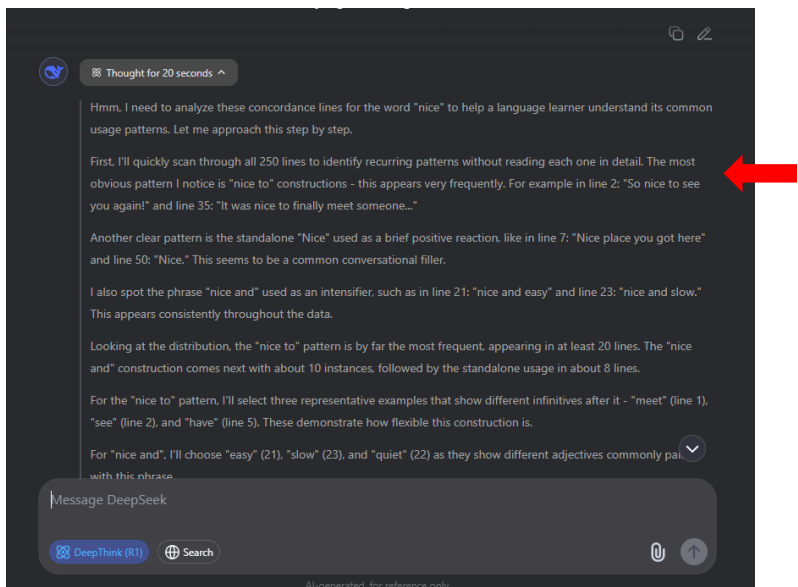
8. Click the Send button



- Wait a few seconds while DeepSeek processes your request. The DeepThink feature will show as "Thinking" to indicate that it is analyzing your input.



You can read through the explanation to understand how the AI arrives at its response. This transparency helps guide prompt refinement and interpretation.



You will receive an in-depth analysis based on the concordance lines.

Based on the provided concordance lines, the most important usage patterns for the word "nice" are:

- 1. "Nice to [verb]..." (Expressing positive feelings about an action or experience):
 - This is the most frequent pattern. It expresses appreciation, pleasure, or politeness about meeting, seeing, talking, having something, etc.
 - Examples:
 - Line 1: "Nice to meet you both."
 - Line 2: "So nice to see you again!"
 - Line 8: "Nice talking to you, Zoltan."
- 2. "Nice [noun]" (Describing something positively):
 - "Nice" directly describes a noun (person, place, thing, idea) in a positive way, meaning pleasant, good, agreeable, or attractive.
 - Examples:
 - Line 7: "Nice place you got here." (Also contains "A nice guy")
 - Line 27: "Nice job with those defenses."
 - Line 30: "Nice Pip-Boy, by the way."

- 3. "Nice and [adjective]" (Emphasizing a desirable quality):
 - This pattern uses "nice and" to intensify the following adjective, emphasizing that something is satisfactorily or pleasingly in that state (e.g., cool, quiet, slow, secure).
 - Examples:
 - Line 21: "I'd hoped to go out in my sleep, nice and easy."
 - Line 22: "Nice and quiet. Shouldn't be any disturbances."
 - Line 23: "Take it nice and slow, no sudden moves..."

Less frequent but notable pattern:

- 4. "Nice..." (Sarcastic/Negative - often "Nice try"):
 - Sometimes "nice" is used sarcastically, especially in the phrase "Nice try," meaning an attempt was unsuccessful or not convincing. It can also stand alone sarcastically.
 - Examples:
 - Line 4: "Nice try, but I just can't let that slide."
 - Line 18: "That's not nice. That's not nice AT ALL." (Implies the opposite)
 - Line 41: "Nice try!"

Summary for Learner:

- Use "Nice to [verb]..." to politely say you enjoy meeting, seeing, or doing something with someone (e.g., "Nice to meet you," "Nice to see you," "Nice talking to you").
- Use "Nice [noun]" to describe something as pleasant, good, or attractive (e.g., "Nice place," "Nice job," "Nice dog").
- Use "Nice and [adjective]" to emphasize something is pleasingly in a certain state (e.g., "Nice and quiet," "Nice and slow," "Nice and cool").
- Be aware that "Nice try" is often used sarcastically when someone fails at something. Sometimes just saying "Nice..." can also sound sarcastic depending on the situation.

Prompt Sample 2: Common patterns, collocation, and colligation identification

Role: You are a linguistics data analysis expert specializing in corpus linguistics and discourse analysis.

Purpose or Goal: The analysis will be used as a guideline to support language learning for me as a learner.

Input: I will provide you with XXX concordance lines for the word X. You do not need to analyze all XXX lines. Instead, focus only on the most important lines that clearly show how the word is commonly used.

Context: This is part of a classroom activity. The concordance lines were extracted from a corpus database using CorpusMate.

Constraints and Guidelines:

- Use only the concordance lines I provide below.
- Do not assume any additional context beyond what appears in the lines.
- Focus only on what is observable in the data.
- Identify frequent and salient patterns of usage.
- Avoid making up examples or interpreting beyond the data given.

Desired Output:

- Identify the word X for 1) the ****most important usage pattern(s)****, 2) ****most important collocation****, and 3) ****most important colligation****.
- Present your findings in a ****simple and structured summary****, explained in ****non-technical terms****.
- Give ****3 real examples**** taken directly from the concordance lines for each pattern.

You will receive an in-depth analysis based on the concordance lines.

Based on the provided concordance lines, here are the key patterns for how "nice" is commonly used:

1. Most Important Usage Pattern(s)

- "Nice to [verb] you": Expressing pleasant social interactions (meeting, seeing, talking).
- "Nice [noun]": Describing something positive (place, job, day, etc.).
- "Nice and [adjective]": Emphasizing a desirable quality (quiet, easy, cool, slow).

2. Most Important Collocation (Word Pairing)

- "Nice place": Used frequently to comment on locations.
- "Nice job": Used to compliment someone's work.
- "Nice day": A common greeting or comment on the weather.

3. Most Important Colligation (Grammar Pattern)

- After "is/was": "Nice" almost always follows a form of "be" (is, was, 's) as the main descriptive word.
- At the start of an exclamation: Often used directly to make a positive comment or greeting.
- Followed by "to" + verb: Used to express that an action is pleasant (e.g., *nice to meet*).

Summary with Examples

1. Usage Pattern: "Nice to [verb] you" (Social Pleasantry)

- Shows polite interaction (meeting, seeing, talking). Very common in greetings/farewells.
- *Example 1:* "Nice to meet you both." (Line 1)
- *Example 2:* "So nice to see you again!" (Line 2)
- *Example 3:* "Nice talking to you, Zoltan." (Line 8)

2. Usage Pattern: "Nice [noun]" (Positive Description)

- Describes a thing (place, job, day, guy, etc.) positively. Very frequent.
- *Example 1:* "Nice place you got here." (Line 7)
- *Example 2:* "Nice job with those defenses." (Line 27)
- *Example 3:* "Nice day, isn't it." (Line 25)

3. Usage Pattern: "Nice and [adjective]" (Emphasizing Quality)

- A fixed phrase highlighting a specific desirable state (quiet, easy, slow).
- *Example 1:* "Nice and quiet. Shouldn't be any disturbances." (Line 22)
- *Example 2:* "I'd hoped to go out in my sleep, nice and easy." (Line 21)
- *Example 3:* "Take it nice and slow, no sudden moves..." (Line 23)

4. Collocation: "Nice place"

- This pairing occurs multiple times, showing it's a common way to comment on locations.
- *Example 1:* "Nice place you got here." (Line 7)
- *Example 2:* "Nice place you have here, citizen." (Line 31)
- *Example 3:* "Nice place. A witcher?!" (Line 233)

5. Collocation: "Nice job"

- Used specifically to praise someone's actions or results.
- *Example 1:* "Nice job with those defenses." (Line 27)
- *Example 2:* "Nice job, ma'am." (Line 17)
- *Example 3:* "Nice job. I guess the Brotherhood will think twice..." (Line 177)

6. Collocation: "Nice day"

- A very common standalone greeting or comment.
- *Example 1:* "Nice day, isn't it." (Line 25)
- *Example 2:* "Nice day." (Line 26)
- *Example 3:* "Have a nice day." (Repeated often, e.g., Lines 102, 105, 108)

7. Colligation: After "is/was"

- "Nice" functions almost exclusively as an adjective following "be" verbs.
- *Example 1:* "This place is a dump. Nice place..." (Line 7 - 'is' implied for 'nice')
- *Example 2:* "That was suspiciously nice of her." (Line 28)
- *Example 3:* "It 's (is) nice to finally have something..." (Line 34)

8. Colligation: Start of an Exclamation/Comment

- Frequently used directly to make a positive remark or greeting.
- *Example 1:* "Nice to meet you both." (Line 1 - Starts the greeting)
- *Example 2:* "Nice work, ma'am." (Line 17 - Starts the compliment)
- *Example 3:* "Nice place you got here." (Line 7 - Starts the comment)

9. Colligation: Followed by "to" + Verb

- Shows that an action (meeting, seeing, talking, having) is pleasant.
- *Example 1:* "Nice to meet you both." (Line 1)
- *Example 2:* "Always nice to have a few extra caps." (Line 5)
- *Example 3:* "Platitudes are nice to hear." (Line 6)