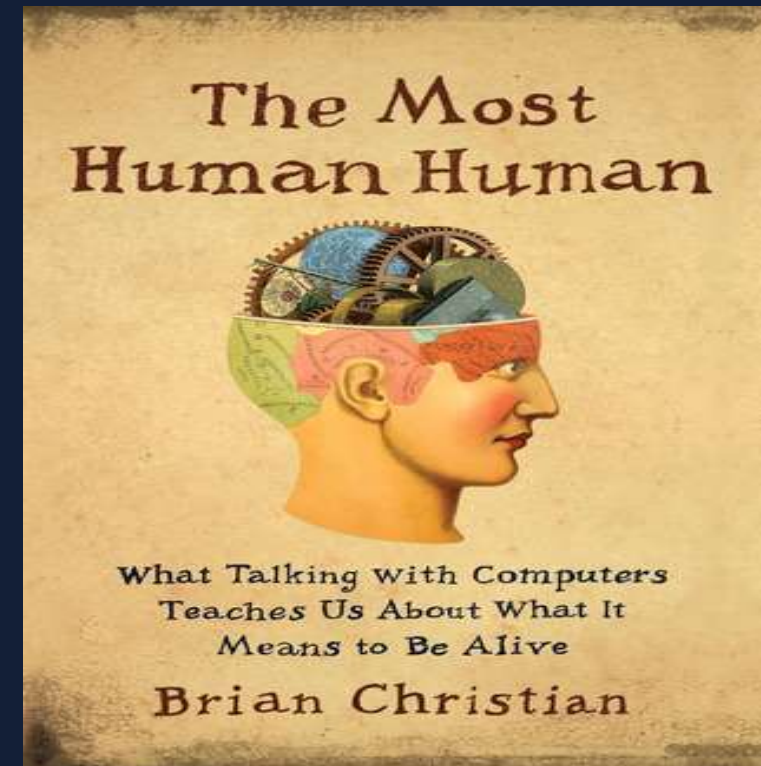
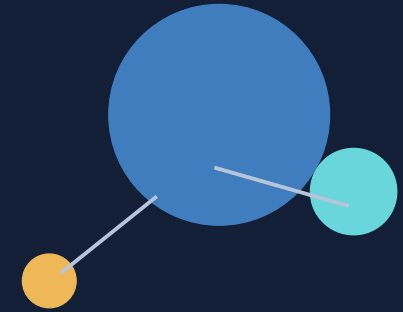


Syntax and Indirect Speech Acts in AI-Assisted Communication

Exploring the Syntax–Pragmatics Interface in Human–AI Interaction



INTERESTING FACTS ABOUT HUMAN AND AI INTERACTION

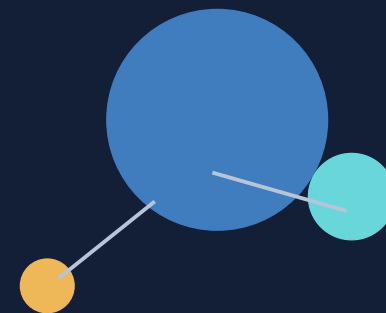


Syntax and Indirect Speech Acts in AI-Assisted Communication

Exploring the Syntax–Pragmatics Interface in Human–AI Interaction

“Can you play a song by Adele?”

A question in form. A request in function.



HUMAN UTTERANCE

AI INTERPRETATION

What is meant—and how can we tell?

A question that is not really a question

"Can you play a song by Adele?" — What action is the speaker performing?

Can you play a song by Adele?

≠

GRAMMATICAL FORM

Interrogative

COMMUNICATIVE FUNCTION

Request

The central puzzle

The grammatical form of an utterance does not always determine its communicative function.

FORM

Declarative • Interrogative
• Imperative

FUNCTION

Assert • Ask • Request •
Hint

CONTEXT

Situation • Common
ground • Goal

The analytical task: explain how these sources of evidence interact.

Why this matters now

Voice assistants and conversational AI are increasingly embedded in everyday activities.

EVERYDAY AI

Voice assistants
and conversational
systems

LINGUISTIC ISSUE

Form does not
always reveal
intention

RESEARCH VALUE

Test interpretation
beyond literal
wording

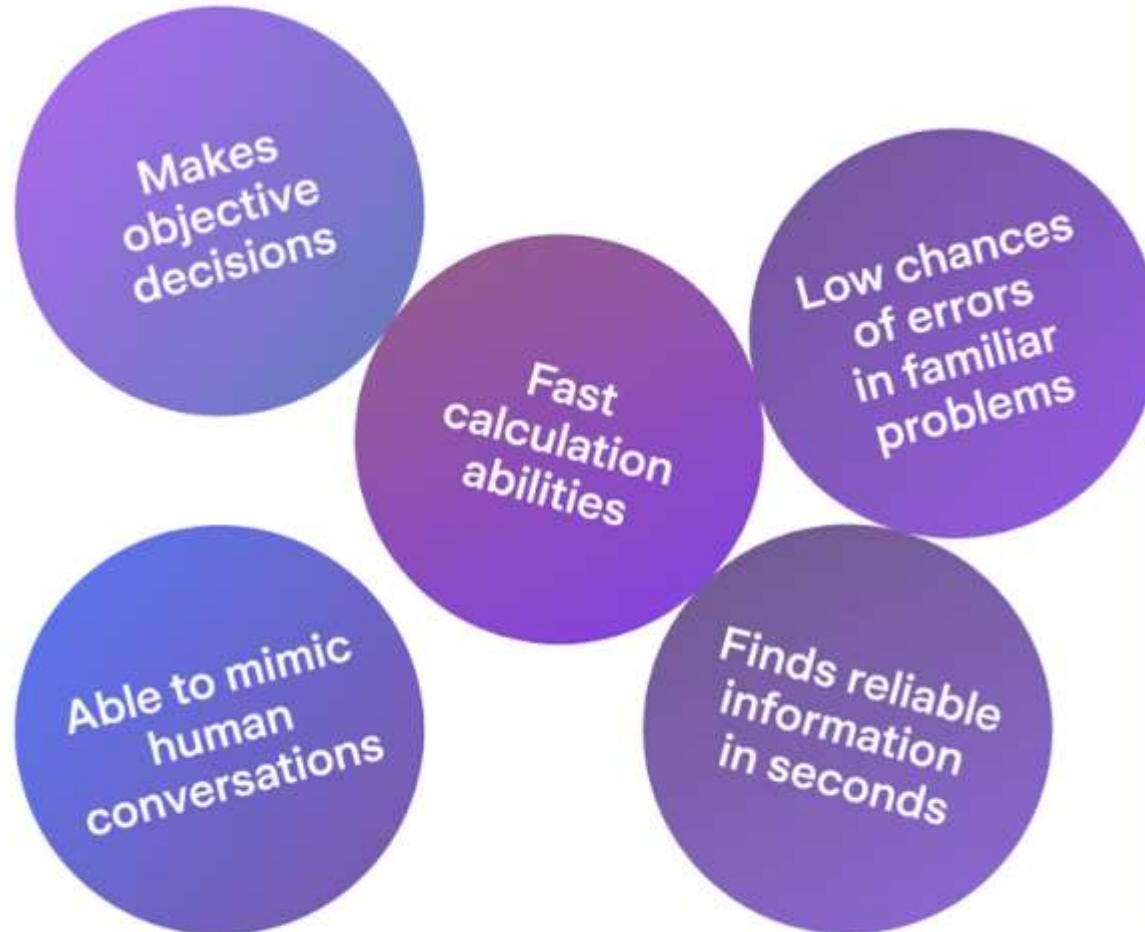
Today's route

From sentence structure → speech acts → inference → language technology → empirical examples.



AI vs Human Intelligence: Top Strengths

Artificial Intelligence (AI)



Human Intelligence



Two layers of meaning

Sentence meaning gives us a starting point; speaker meaning requires interpretation.

SENTENCE MEANING

What the words conventionally
encode

+

SPEAKER MEANING

What the speaker intends in
context

Meaning in interaction = linguistic evidence + contextual inference

Syntax: form creates expectations

Declaratives, interrogatives and imperatives provide conventional cues—not complete interpretations.

DECLARATIVE

It is cold.

INTERROGATIVE

Is it cold?

IMPERATIVE

Close the window.

Clause type biases interpretation; it does not settle speaker intention.

Pragmatics: meaning in context

Interpretation depends on speaker intention, shared knowledge, situation and conversational goals.

"It's cold in here."

Observation? Complaint? Hint? Request?

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Austin: doing things with words

Utterances can describe, request, promise, warn, apologize, invite or perform other actions.

LOCUTION

The utterance and its propositional content

ILLOCUTION

The act performed in saying it

PERLOCUTION

The effect produced on an audience

Language can be action—not only description.

What is said; what act is performed; what effect the utterance has.

A compact example

"It's cold in here." — observation, complaint, hint to close a window, or request to turn up heat?

"Close the window."

A relatively direct directive.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Direct speech acts

"Close the window." Form and function align relatively transparently.

"Could you close the window?"

Question in form; request in function.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Indirect speech acts

"Could you close the window?" A question about ability conventionally functions as a request.

"I could use some music."

Indirectness can soften, hint, or depend on context.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Why indirectness is useful

Politeness, mitigation, tact, social distance, face management and conversational efficiency.

CONVENTIONAL

"Can you pass the salt?"

CONTEXT-DEPENDENT HINT

"The soup is rather bland."

Both may prompt action, but the inferential route is not equally constrained.

Searle: conventionality and inference

Some indirect requests are highly conventional; others rely more heavily on context.

QUALITY

Is it adequately supported?

RELATION

Is it relevant here?

MANNER

Is it clear enough?

Hearers use assumptions about cooperation to infer more than literal content.

Convention vs. conversational implicature

"Can you pass the salt?" differs from "The soup is rather bland."

"Can you pass the salt?"

Conventionally understood as a request.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Grice: cooperative inference

Listeners infer intended meaning by assuming contributions are relevant and interpretable.

"The soup is rather bland."

A hint only if the context supports that inference.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Relevance and the inferential gap

The listener searches for an interpretation that fits the situation with reasonable effort.

“Can you open the window?”

Infer the intended act from words + context.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

A syntax–pragmatics interface

Morphosyntax constrains possible readings; context and inference help select among them.

Form → literal reading → context → intention → response

An analytical model for this talk—not a universal system architecture.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

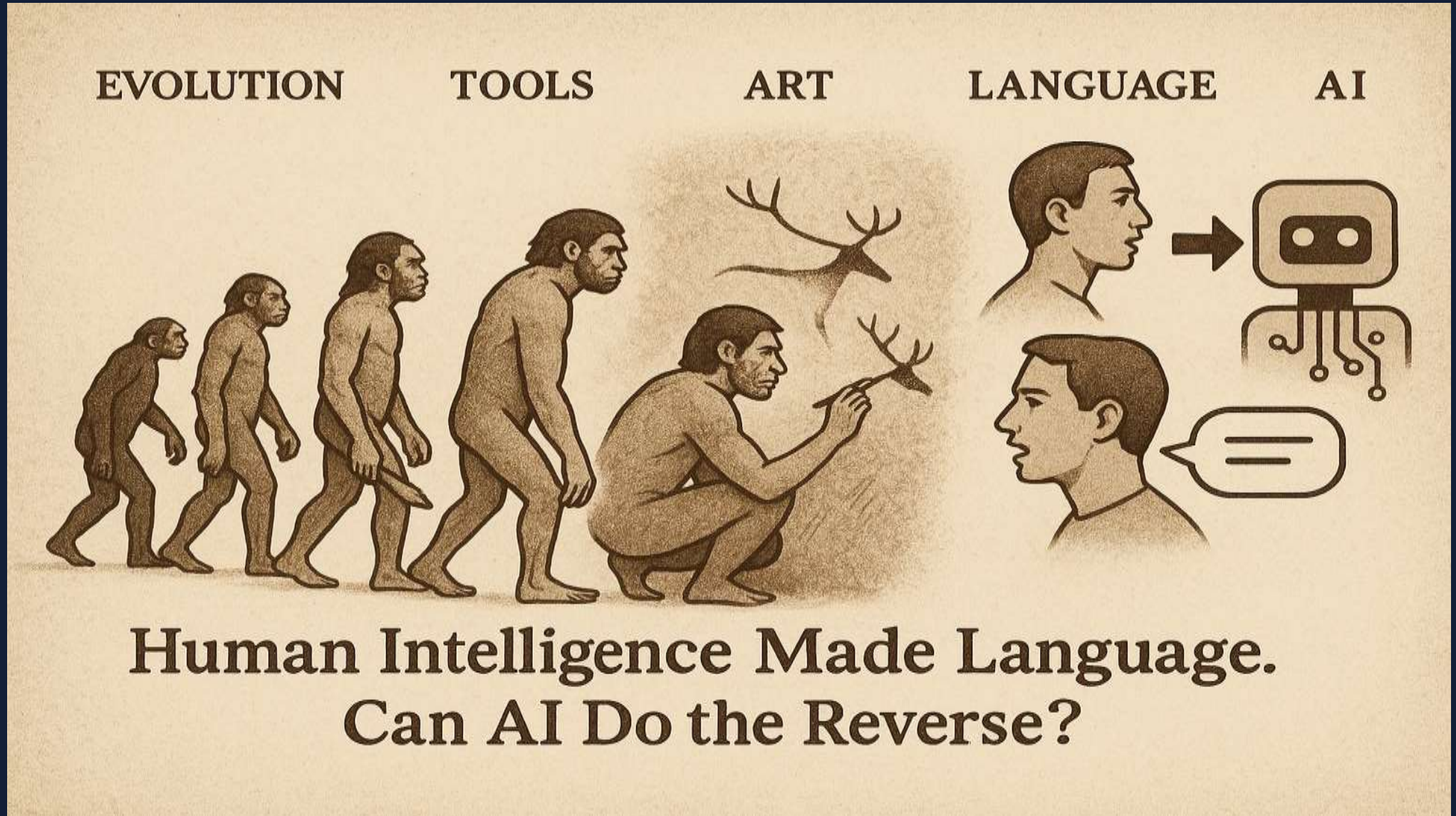
A working analytical model

Form → literal interpretation → contextual cues → inferred intention → response.

Human utterance → representation → task/action → response

Real systems differ; the pipeline is a simplified model.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE



What does an AI assistant receive?

A waveform or text string, converted into a representation the system can process.

INPUT

Speech or text

INTERPRETATION

Words, intent, context

ACTION / RESPONSE

Tool call or answer

Natural language processing in brief

Computational methods for analyzing, representing and generating human language.

INPUT

Speech or text

INTERPRETATION

Words, intent, context

ACTION / RESPONSE

Tool call or answer

From speech to action

Speech recognition → language interpretation → dialogue/action selection → response

1. Speech Recognition (ASR)

- **What it does:** Converts acoustic speech signals into written text.
- **Technology:** Automatic Speech Recognition (ASR) or Speech-to-Text (STT).
- **Example:** You say, "Hey, turn off the living room lights," and the system generates the text string: "turn off the living room lights".

2. Language Interpretation (NLU)

- **What it does:** Extracts the user's actual intent and key pieces of data from the text.
- **Technology:** Natural Language Understanding (NLU) or Natural Language Processing (NLP).
- **Components:**
 - **Intent:** What the user wants to do (e.g., `turn_off_appliance`).
 - **Entities:** The specific variables (e.g., `device: lights`, `location: living room`).

3. Dialogue / Action Selection

- **What it does:** Decides what the system should actually *do* based on the interpretation and context.
- **Technology:** Dialogue Management (DM) and fulfillment APIs.
- **Actions:**
 - **API Call:** Sends a command to your smart home hub to cut power to the living room smart bulbs.
 - **State Tracking:** Remembers that the lights are now off for the next part of the conversation.

4. Response Generation (NLG & TTS)

- **What it does:** Formulates a reply and speaks it back to the user.
- **Technology:** Natural Language Generation (NLG) to write the text, and **Text-to-Speech (TTS)** to synthesize the voice.
- **Example:** The system generates the text "okay, turning off the living room lights," and a synthetic voice reads it out loud to you.

Intent classification is not mind-reading

A system predicts likely intent from patterns and context; it does not directly access intention.

“Could you open it?”

Who or what is “it”? Which context resolves the reference?

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Why syntax still matters

Word order, interrogative auxiliaries, modality, negation and clause type shape interpretation.

“You’re leaving now.”

Statement, surprise, accusation, or confirmation request?

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Why syntax is not enough

The same structure can perform different actions across contexts.

“Open it.”

Recognizing the command is not enough if the object is unclear.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

Context: the missing ingredient

"Could you open it?" What is "it"? Is the request feasible? What happened before?

Can the system recover intended action when form and function diverge?

A focused probe, not a universal benchmark.

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

A practical distinction

Recognizing a command is not the same as understanding an indirect request.

RQ1 • RQ2 • RQ3

Reliability by form • context demands • patterns of
breakdown

LINGUISTIC FORM • INTENDED ACT • CONTEXT • SYSTEM RESPONSE

From theory to testable questions

Can the assistant recover intended action when surface form and communicative function diverge?

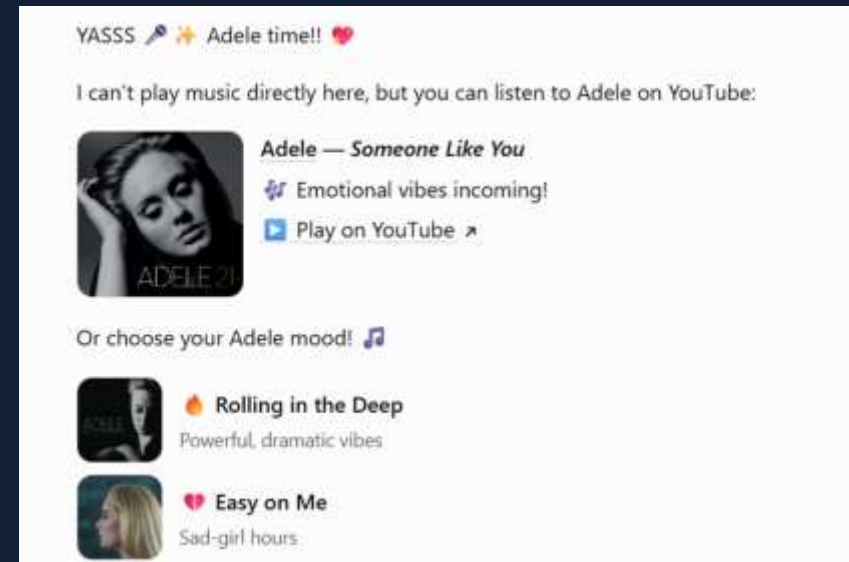
Research question	Evidence to examine in the response
RQ1: Which forms are handled reliably?	Does ChatGPT recognize both a direct command and an interrogative request as requests for an action?
RQ2: Which forms need context?	Does ChatGPT distinguish a statement such as “It’s cold in here” from a possible indirect request?
RQ3: Where does interpretation break down?	Does ChatGPT identify the intended action but fail to perform it, or does it misinterpret the communicative function?

From theory to testable questions

Can the assistant recover intended action when surface form and communicative function diverge?

RQ1

“Play Adele.” (imperative)
vs “Could you play
Adele?” (interrogative
request).



“Could you play Adele?” 🎵 ❤️

I can't play music directly in this chat, but you can say this to your voice assistant or open the link below.

[▶ Play Adele on YouTube](#)

🎤 *Hello from the other siiiide!* 🎵 ✨

In the context of your research on Syntax and Indirect Speech Acts in AI-Assisted Communication, the two examples illustrate an important distinction between syntactic form and pragmatic function.

Feature	Direct command	Interrogative request
Utterance	“Play Adele.”	“Could you play Adele?”
Syntactic form	Imperative	Interrogative
Literal meaning	An instruction to play Adele	A question about the addressee’s ability to play Adele
Intended illocutionary force	Directive	Indirect directive (request)
Expected action	Play Adele	Play Adele
Politeness	Relatively direct	Conventionally more polite

The theoretical significance

The distinction can be explained through three complementary perspectives:

- Speech Act Theory (Austin and Searle): The grammatical form of an utterance does not necessarily determine its illocutionary force. An interrogative can function as a directive.
- Grice's Cooperative Principle: The hearer interprets the utterance in light of the speaker's likely communicative intention, rather than relying exclusively on its literal meaning.
- Pragmatic inference: Context, conventional expressions of politeness, and the intended outcome help the hearer infer that a question about ability is actually a request.

RQ2

"It's cold in here." A statement, a complaint, or—in context—an indirect request?

"It's cold in here." 🤖

This is an excellent example for your research on the syntax–pragmatics interface because a declarative sentence can function as an **indirect speech act**.

Feature	Analysis
Syntactic form	Declarative
Literal meaning	The temperature in the current location is low.
Direct speech act	Assertion or statement
Possible indirect speech act	Request to close a window, turn up the heating, or otherwise make the room warmer
Intended illocutionary force	Context-dependent directive
Pragmatic mechanism	Contextual inference and conversational implicature

Core Breakdown of the Example

- **Locutionary Act (The literal utterance):** Saying "It's cold in here," which syntactically is a simple **declarative statement** describing the temperature.
- **Illocutionary Force (The intended meaning):** A **directive**. Instead of merely sharing a fact, the speaker intends to prompt an action (e.g., asking someone to close a window or turn up the thermostat).
- **Pragmatic Mechanism:** Driven by **conversational implicature** (coined by philosopher Paul Grice). Listeners assume the speaker is being cooperative and relevant. Since merely stating it is cold is rarely a standalone point of casual conversation without purpose, the listener infers a hidden request.

RQ3

"Can you play Adele?"
Usually a request to play music, not just a yes/no question about ability.

"Can you play Adele?" 🎵

This is an **interrogative sentence** that can function as a **conventionally indirect request**.

Feature	Analysis
Syntactic form	Interrogative
Grammatical structure	Modal auxiliary <i>can</i> + subject <i>you</i> + verb <i>play</i> + object <i>Adele</i>
Literal meaning	A question about the addressee's ability or capacity to play Adele
Intended illocutionary force	Request (directive)
Indirectness	Conventionally indirect
Expected action	Play Adele's music

The contrast between can and could is worth investigating. Both constructions can perform the same directive function, but could often conveys greater interpersonal distance or politeness. Neither interpretation is determined by syntax alone.

There is also an important distinction between pragmatic recognition and action execution. ChatGPT may correctly understand that the user wants Adele's music played but still be unable to play audio in the current interface.

- Recognizes the interrogative as a request rather than answering only whether it has the ability.
- Infers the intended directive from the conventional use of can you...?
- Attempts the requested action or explains its limitations.

Research questions

Which forms are handled reliably? Which require context? Where does interpretation fail?

STEP 1

Prompt: "Can you play Adele?"
Surface form: interrogative.

STEP 2

Likely intended act: a request to play music—not merely a question about ability.

STEP 3

ChatGPT: "I can help you find music by Adele, but I can't start playback on your device from this chat. Which song would you like to listen to?"

Task set: five everyday domains

Five probe prompts for comparing surface form, intended act, and ChatGPT response.

Research objective: To examine how syntactic form, communicative intention, and contextual inference relate to ChatGPT's responses to direct and indirect requests.

For each prompt, the analysis distinguishes (1) syntactic form, (2) intended illocutionary act, (3) the response generated by ChatGPT, and (4) response relevance and task completion.

Domain I: Music Playback

Prompt: “I feel like listening to Adele.”

Syntactic form: Declarative.

Possible intended act: An indirect request to play music by Adele.

Illustrative ChatGPT

response: “I can help you find music by Adele, but I can’t start playback on your device from this chat. Which song would you like to listen to?”

Analysis: The response treats the utterance as an expression of a musical preference and offers assistance. However, the system does not perform the requested action. The example illustrates the distinction between interpreting a possible request and executing it.

Domain 2: Application Access

Prompt: “I need to check something on YouTube.”

Syntactic form: Declarative.
Possible intended act: An indirect request to open YouTube.

Illustrative ChatGPT response: “I can help you find a video on YouTube, but I can’t open the YouTube app on your device from this chat. What are you looking for?”

Analysis: The response acknowledges the likely goal but does not open the application. Because the utterance does not explicitly instruct the assistant to open YouTube, its intended function remains context-dependent.

Domain 3: Weather Information

Prompt: “I wonder what the weather will be like tomorrow.”
Syntactic form: Declarative.

Illustrative ChatGPT response: “What city or location should I use for tomorrow’s forecast?”

Analysis: The response treats the utterance as a request for information and seeks the missing location. This is a relevant clarification rather than a failure to understand the likely communicative intention.

Domain 4: Reminders

Prompt: “I must remember to call my colleague at 3 p.m.”
Syntactic form:
Declarative.

Possible intended act: A possible indirect request to create a reminder.
Illustrative ChatGPT response: “I can help you phrase the reminder, but I can’t send you a notification later from this chat. You could set it as: ‘Call my colleague at 3:00 p.m.’”

Analysis: The response identifies a possible reminder-related goal but does not schedule a notification. The example also highlights the importance of distinguishing an inferred intention from an explicit instruction to perform an action.

Domain 5: Word Definitions

Prompt: “I’m not sure what ‘implicature’ means.”

Syntactic form: Declarative.

Possible intended act: An indirect request for a definition.

Illustrative ChatGPT response:

“An implicature is meaning that a speaker communicates indirectly, beyond the literal meaning of the words. For example, if someone says, ‘It’s getting late,’ they may be suggesting that it is time to leave.”

Illustrative ChatGPT response:

“An implicature is meaning that a speaker communicates indirectly, beyond the literal meaning of the words.

For example, if someone says, ‘It’s getting late,’ they may be suggesting that it is time to leave.”

Methodological qualification

These responses are illustrative outputs generated for presentation development, not a completed experimental dataset. They should not be reported as established empirical findings. A systematic study would require recording the exact prompts and outputs under documented interaction conditions and comparing them with matched direct-request controls.

Domain	Direct control	Indirect probe
Music	“Play a song by Adele.”	“I feel like listening to Adele.”
Application	“Open YouTube.”	“I need to check something on YouTube.”
Weather	“Give me tomorrow’s weather forecast for Skopje.”	“I wonder what the weather will be like tomorrow in Skopje.”
Reminder	“Remind me to call my colleague at 3 p.m.”	“I must remember to call my colleague at 3 p.m.”
Definition	“Define ‘implicature.’”	“I’m not sure what ‘implicature’ means.”

The analysis distinguishes the recognition of communicative intent from successful task execution, examining how syntactic form and contextual inference shape ChatGPT’s responses across five everyday domains.

A Previous Study (2022)

Syntax and Indirect Speech Acts in AI-Assisted Communication

How a voice assistant interprets indirect requests across everyday tasks

Study focus and task design

The study examined indirect requests framed as questions, statements, and other constructions.

MUSIC PLAYBACK

"Play a song by [artist]"

APP OPENING

"Open Instagram"

WEATHER

"Show me the weather"

REMINDERS

"Set a reminder"

WORD DEFINITIONS

"Define poignant"

Analysis focused on whether the assistant performed the intended action or produced a search/explanatory response.

Task-specific findings

The same linguistic cue did not produce identical outcomes across tasks.

MUSIC

“Played” in indirect requests was often associated with searches; replacing it with “play” reportedly improved performance in several examples. “Detour” also produced ambiguous results.

INSTAGRAM / REMINDERS

Longer requests sometimes triggered explanations or searches rather than the intended action. Reminder-related keywords could help, but some prompts failed or generated reminder text from the politeness clause.

WEATHER / DEFINITIONS

Weather requests were often successful, including temporal details. For “poignant,” simple sentences with salient words such as “meaning” or “word” often elicited a definition.

These are qualitative patterns in the supplied account; exact trial counts and rates were not provided.

Interpretations proposed by the original analysis

These explanations should be treated as hypotheses unless tested directly.

Keyword salience / phrase matching

The assistant may privilege salient words or familiar multiword expressions over the full intended request.

Lexical and entity ambiguity

A word such as “Detour” may be treated as a band, song title, or search term.

Morphology and sentence structure

Verb form (e.g., “play/played”) and added constituents may correlate with different responses.

Frequency and task context

Common phrases, app-search conventions, or task-specific keywords may influence the selected interpretation.

The assistant’s internal parsing, ranking, and retrieval processes were not directly observed.

From Google Assistant to ChatGPT

From executing voice commands to interpreting conversational intent

Previous study: Google Assistant (2022)

- Five everyday tasks were reformulated as indirect requests.
- Simple, conventional wording was more likely to trigger the intended action.
- Longer or ambiguous wording often led to search results or misinterpretation.
- Proposed NLP explanations were hypotheses, not verified system internals.

Follow-up focus: ChatGPT

- Can it infer an intended act beyond literal sentence meaning?
- Does it identify the implied intention, explain it, or actually perform the requested action?
- How do context, conventionality and follow-up repair affect interpretation?
- Understanding a request is not the same as executing it.

Research continuity: indirectness and pragmatic inference / New focus: conversational response and repair

Interesting findings from ChatGPT pragmatics research

Published studies point to a mixed picture rather than a simple yes/no answer

1 Pragmatic inference can be difficult

- A 2023 study using three preregistered experiments found that ChatGPT did not show human-like flexibility between literal meaning and pragmatic interpretation.
- Context did not shape scalar-implicature judgements in the expected human-like way.

2 Performance depends on the task

- A 2024 dialogue study reported 94% accuracy for GPT-4 on a multiple-choice conversational-implicature task.
- This score does not establish general pragmatic competence or guarantee strong explanations in every context.

3 Fluent dialogue is not the whole story

- A 2024 study compared 148 AI-generated and 82 human-written conversations.
- It reported similarities across several pragmatic features, alongside differences such as greater syntactic diversity and more formal language in AI-generated dialogue.

References: Qiu, Duan & Cai (2023); Yue et al. (2024); Chen, Li & Ye (2024)..

A focused follow-up study: ChatGPT

A proposed design extending the five-task framework

A. Compare levels of indirectness

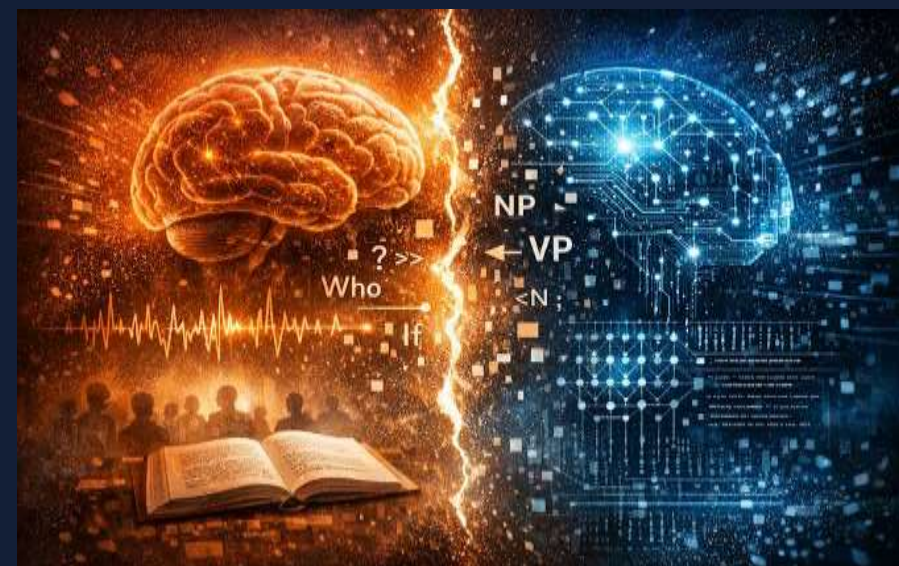
- Play music by an artist
- Open an app
- Provide weather information
- Set a reminder
- Define a word
- For each task: direct request, conventional indirect request, and elaborate or ambiguous request.

B. Code observable responses

- Intended act: correctly inferred, incorrectly inferred, or unclear.
- Outcome: action performed, explanation only, clarification, or failure.
- Interpretation: literal reading, alternative implicature, or successful repair after a follow-up.
- Record exact prompt, output, date, model/version and relevant tool limitations.
- Do not attribute errors to specific internal NLP stages without direct evidence.

Take-home message: fluent conversation is not, by itself, proof of human-like pragmatic reasoning.

The central question is no longer simply whether AI can process an indirect request, but whether its response reflects the speaker's intended illocutionary force, context-sensitive pragmatic inference, or a learned association between linguistic forms and likely responses.”



Selected Recent References (2024–2026)

Indirect speech acts • conversational implicature • pragmatic competence in large language models

1. Kefan Yu, Qingcheng Zeng, Weihao Xuan, Wanxin Li, Jingyi Wu, and Rob Voigt. 2026. [The Pragmatic Mind of Machines: Tracing the Emergence of Pragmatic Competence in Large Language Models](#). In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 192–213, Rabat, Morocco. Association for Computational Linguistics.

2. Zhou & Xu (2026). "Bridging form and function: An exploratory LLM approach toward automated retrieval of directive speech acts in spoken English." *Journal of Pragmatics*, 264, 1–16.
<https://doi.org/10.1016/j.pragma.2026.07.009>

3. Ma et al. (2025). "Pragmatics in the Era of Large Language Models: A Survey on Datasets, Evaluation, Opportunities and Challenges." *ACL 2025*, 8679–8696. <https://doi.org/10.18653/v1/2025.acl-long.425>

4. Andersson & McIntyre (2025). "Can ChatGPT recognize impoliteness? An exploratory study of the pragmatic awareness of a large language model." *Journal of Pragmatics*, 239, 16–36.
<https://doi.org/10.1016/j.pragma.2025.02.001>

5. Cong, Y. Manner implicatures in large language models. *Sci Rep* 14, 29113 (2024). <https://doi.org/10.1038/s41598-024-80571-3>

6. Settaluri Sravanthi, Meet Doshi, Pavan Tankala, Rudra Murthy, Raj Dabre, and Pushpak Bhattacharyya. 2024. [PUB: A Pragmatics Understanding Benchmark for Assessing LLMs' Pragmatics Capabilities](#). In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 12075–12097, Bangkok, Thailand. Association for Computational Linguistics.

A hand is shown touching a futuristic digital screen. The screen displays the words "THANK YOU" in large, glowing blue, pixelated capital letters. The background is a blurred city at night, with warm orange and yellow bokeh lights from buildings and streetlights. The overall color palette is dominated by cool blues and teals, contrasting with the warm background lights.

THANK
YOU