

Syntax vs Semantics

Samet Efe Keskin, PHIL308 Interaction Paper

May 7, 2026

Abstract

This paper examines how Symbol Manipulation & Understanding is linked and whether a perfect context-sensitive input-output machine can achieve full understanding. The paper draws on previous literature and the theories of Functionalism, Biological Naturalism, and Dualism. I focus on Searle's *Chinese Room Argument*, and discuss the analogy on both opposing sides. I follow an AI interaction protocol on the matter of the question, then break down the arguments and positions. In addition to observations based on the protocol, I use relative studies to conclude that understanding requires meaning and experience, which cannot be produced by syntax alone.

Keywords: The Chinese Room Argument, Functionalism, symbol-grounding, Computation, Intentionality

1 Motivation and the Question

The discussion of the question includes two key terms that to be defined formally. The term *syntax* is defined to be the formalization of syntactic properties of language [Tallerman, 2020], and *semantics* refers to the meaning behind language, and significance in whole [Haugeland, 1980]. The main question of whether genuine understanding can be captured by pure syntax is introduced in Descartes' dualism in the most radical form. By characterizing cognition and substance differently, Descartes argues that understanding can not be reduced to system manipulation. The recorded answers and objections from Behaviorists and Functionalists followed the discussion. Later, Empiricists such as David Hume and John Locke asserted that ideas come from experience. Their arguments implied that pure logic, understood as the formal study of patterns, cannot capture sensory information. Therefore, if symbols don't include experience, how can they mean anything? Like the other scholars, I try to provide an answer to the question. Still, I'll use an AI interaction as input and assess the observations against my hypothesis.

2 Interaction and Analysis of the Protocol

As stated in the task description, I started with the mandatory prompts and asked whether a computer can actually understand a language through syntax. The initial response from ChatGPT emphasized the core claim that the lack of semantics is why a rule-based system cannot genuinely understand. As we discussed in class, he provided the Chinese Room example to illustrate his position.

2.1 The Chinese Room Example

In his 1984 book [Searle, 1980], John R. Searle introduced a thought experiment about an imaginary person in a room who doesn't speak Chinese. They receive Chinese characters and produce appropriate responses in Chinese using an instruction manual. Searle argued that what was happening was not a genuine understanding but a purely syntactic formation. He added that passing behavioral tests does not constitute understanding. In my dialog, ChatGPT provided the experiment along with objections and their answers.

- **Objection 1:** Although the person doesn't understand Chinese itself, the system holistically does.
- **Response 1:** *"Even in system level, still every iteration is syntactic."*
- **Objection 2:** If the system indeed mimics Chinese perfectly, isn't that enough?
- **Response 2:** *"Behavior alone is insufficient."*

One clear alignment in the responses reflects Churchland's objection to the idea that no understanding is being produced in the room [Churchland and Churchland, 1990]. He argued that this formal rule-based manipulation could enter a robot's body and use sensory information as input. Although even in that case the "understanding" is limited and insufficient, the author believed that "the Brain Simulator Reply" concept is enough to reject the idea that there is no understanding at all. In the second prompt, when I was asked to pick the strongest objection and defend the stated thesis, AI actually affirmed the objection and suggested that Searle might be advancing the wrong argument. Just as Churchland, ChatGPT states that one cannot disprove the existence of an understanding at a holistic level, simply because of insufficiency. In addition to its answer, the system also expands on the Functionalist idea. It reduces the objection to a dilemma, forcing one to choose between them to defend its initial position.

1. Understanding is indeed equivalent to functional capacity, and if the system mimics the understanding it does, since it can.
2. If understanding is beyond function, then one is entitled to explain what it is, how it is detectable, and whether it matters.

It is also evident that the first side of the dilemma aligns with Fodor's argument that syntax has a semantic interpretation. He argued that symbols have semantic value because we interpret them semantically [Fodor, 1987]. If syntax can manipulate the representations of the outer world, then it is already part of understanding. Once I asked for the possible counterexamples and cases supporting its initial position, it provided the case of the *Perfect Phrasebook Tourist*.

The main objection was that although the outputs were indistinguishable, it is unclear who really knows Chinese, and it is just a behavioral test. When nuanced and inter-layered conversations are tested dynamically, the experiment cannot guarantee that the tourist will again pass. ChatGPT concluded that near-zero semantic understanding may still exist with perfect language use. After the last prompt, I asked a follow-up on the tourist analogy. If the tourist, and its archetype, is fully capable of producing a response that best fits the situation, do we have any tools to distinguish with real understanding from the alternative one? The response I received was mildly surprising, as it offered a perspective on why understanding matters. ChatGPT replied that although we cannot distinguish behavioral understanding, having a theory grounded in the following conditions would still provide a benchmark for the philosophical aspect of the discussion.

- If a system is grounded in perception;
- and has an internal state that is genuine on subjects mentioned, then it understands.

I can only interpret the second part of the conditions as having a semantic content over the words and their meanings.

3 Conclusion

Hence, when I asked for the final position in the discussion, AI provided a series of premises that show a strong correlation with Biological Naturalism and, therefore, a major connection with Searle. The prompt responses I had included explicit definitions on semantics not being the same as syntax, and the idea that meaning could only come from semantics. Although it provided the same answer, it didn't offer answers close to Descartes' Dualism, as it didn't address the soul or its link to the non-material world. Therefore, the AI's overall stance in this discussion is perfectly aligned with Searle and Biological Naturalism, emphasizing that a genuine understanding must consist of semantic meaning, which cannot come from syntax.

References

[Churchland and Churchland, 1990] Churchland, P. M. and Churchland, P. S. (1990). Could a machine think? *Scientific American*, 262(1):32–37.

[Fodor, 1987] Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. The MIT Press, Cambridge, MA.

- [Haugeland, 1980] Haugeland, J. (1980). Programs, tips, and brains. *Behavioral and Brain Sciences*, 3(3):432–433.
- [Searle, 1980] Searle, J. R. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences*, 3:417–457.
- [Tallerman, 2020] Tallerman, M. (2020). *Understanding Syntax*. Routledge, New York, NY, 5th edition.

A Appendix

Hi, I am working on a Philosophy of Mind assignment. I will ask you several questions. Your task is to act as a good philosophical interlocutor: choose a reasonable position on the topic, defend it consistently, and respond to objections as strongly as you can. Please do NOT stay neutral—take a stance and justify it. Here is my first question..

I'm ready. Ask your first question, and I'll take a clear position and defend it consistently. *Can a computer understand a language merely by manipulating symbols according to rules? Please take a clear position and defend it.* I'll take a clear stance: No—mere symbol manipulation is not sufficient for genuine understanding of a language. It can simulate understanding impressively, but it does not constitute understanding.

1. Core Claim

A system that only manipulates symbols according to formal rules (syntax) lacks semantics—it does not grasp meaning. Understanding requires more than rule-following; it requires a connection between symbols and what they represent.

2. The Strongest Intuition: The Chinese Room

Imagine a person inside a room who does not know Chinese. They receive Chinese characters, follow an instruction manual (in English), and output appropriate Chinese responses.

From the outside, it looks like the room understands Chinese.

But internally:

The person is just following rules There is no comprehension of meaning It's purely syntactic manipulation

Conclusion: Passing behavioral tests (like conversation) does not imply understanding.

3. Why Syntax Semantics

Symbol manipulation systems operate on:

Form (syntax): shapes, patterns, rules But understanding requires: Meaning (semantics): what symbols refer to

Example:

A system can process the sentence “Fire is hot” But does it know what heat feels like? What fire is?

Without grounding, symbols are:

just tokens related to other tokens—an empty loop

This is known as the symbol grounding problem.

4. Strong Objection: “But if behavior is identical, isn’t that enough?”

This is the classic functionalist reply:

If a system behaves exactly like a human, then it understands.

My response:

Behavior alone is insufficient.

Why?

You can simulate a process without instantiating it A simulation of a hurricane is not wet A simulation of digestion does not digest Similarly, simulating understanding is not the same as understanding

So:

Perfect linguistic performance genuine comprehension

5. Stronger Objection: “The system as a whole understands”

Some argue:

The person doesn’t understand Chinese, but the whole system (person + rules + room) does.

My response:

Even at the system level:

Every step is still syntactic No component introduces meaning Combining non-understanding parts does not magically create understanding

There is still:

no point where symbols become meaningful

6. What Would Be Required for Understanding?

For real understanding, a system needs:

Semantic grounding → connection to the world (perception, action, embodiment) Intentionality → aboutness (thoughts referring to things) Possibly consciousness → subjective experience may be necessary for meaning

This suggests:

Understanding is not just computation—it is embedded cognition

7. Final Position

A computer that only manipulates symbols according to rules:

Can simulate conversation Can pass tests like the Turing Test But does not genuinely understand language

Because:

Syntax alone cannot generate semantics.

If you want, push back—e.g., defend functionalism, AI understanding, or attack the Chinese Room—and I'll respond rigorously.

Give me the strongest objection to your view—for example, from Searle's Chinese Room argument. Good—this is exactly the pressure point where my position is most vulnerable.

I'll give you the strongest possible objection—one that, if successful, undermines my claim entirely.

The Strongest Objection: The Systems Reply (Against the Chinese Room)

The objection is this:

Searle is attacking the wrong target. It's true that the person in the room doesn't understand Chinese—but the whole system does.

1. The Core Argument

In the Chinese Room:

The human = one component The rulebook = another The symbol database + input/output process = the rest

Together, they form a complete cognitive system.

The objection says:

You cannot deny understanding at the system level just because one part lacks it.

2. Why This Is Powerful

Because we already accept this structure elsewhere:

A single neuron doesn't understand English A single transistor doesn't understand anything

But:

The whole brain does.

So the critic presses:

Why should we expect any individual part of a system to understand, if understanding is an emergent property of the whole?

3. Turning My Own Argument Against Me

I argued:

“All parts are just manipulating symbols, so no understanding exists.”

The objection replies:

That's like saying: "Each neuron just fires electrically, so the brain doesn't think."

This is a fallacy of composition:

Just because parts lack a property Doesn't mean the whole lacks it 4. The Internalization Thought Experiment (Stronger Version)

Now the objection gets sharper.

Suppose:

The person memorizes the entire rulebook Internalizes all symbol manipulations Becomes the whole system

From the outside:

They behave exactly like a fluent Chinese speaker

Now ask:

At what point do we deny they understand Chinese?

If you still say:

"They don't understand"

Then the critic says:

You are ignoring all possible behavioral and functional evidence of understanding.

This leads to a troubling conclusion:

Your notion of "understanding" becomes mysterious, untestable, and possibly meaningless

5. The Functionalist Challenge

The objection ultimately forces this dilemma:

Either:

Option A:

Understanding = functional capacity → If a system behaves like it understands, it does understand

Option B:

Understanding = something beyond function → Then you must explain:

What it is How we detect it Why it matters

If you can't do that clearly, the objection claims:

Your position relies on an undefined "ghost property"

6. Why This Objection Is So Dangerous

Because it pushes my view into a corner:

If I deny the system understands → I risk mysticism If I accept it does → my original claim collapses

It essentially says:

There is no principled difference between "simulating understanding" and "having understanding"

7. Where This Leaves My Position

To survive this objection, I must defend:

A non-functional account of understanding Likely involving: semantic grounding embodiment or consciousness

But that is much harder to justify rigorously.

If you want, next step: you can attack this objection, or ask me to defend my original position against it (which is where it gets really interesting).

Give me a counterexample or hypothetical case that supports your position over the opposing one.

Here's a counterexample that directly targets the Systems Reply and supports my position that symbol manipulation alone is not sufficient for understanding:

The "Perfect Phrasebook Tourist"

Imagine this case:

A tourist visits China with zero knowledge of Chinese. However, they carry:

A perfect phrasebook Indexed for every possible situation With precise instructions like: "If you see this symbol, respond with that symbol"

Over time, they become incredibly fast:

They respond fluently They pass for a native speaker in conversation They never make mistakes

From the outside:

They are indistinguishable from someone who understands Chinese.

But Here's the Key Point

Internally:

They still do not know what any word means They are just matching patterns to responses
No symbol is connected to a concept

If you ask:

"What does this sentence mean?"

They cannot answer—only transform it into another string.

Why This Supports My Position

This case shows:

Perfect behavioral performance is compatible with total lack of understanding

So:

Functional equivalence semantic understanding Input-output success meaning

This directly challenges the Systems Reply:

Even if the whole system behaves perfectly, there may still be no understanding anywhere.

Strengthening the Case

Now make it even stronger:

Suppose:

The tourist memorizes the entire phrasebook No external book remains All processing happens "inside their head"

Now they are:

Fast, fluent, indistinguishable from a native speaker

Yet still:

They are just executing internalized rules without grasping meaning.

The Crucial Insight

This creates a powerful intuition:

There is a gap between "using language correctly" and "understanding language"

And that gap cannot be closed by:

More rules More speed More complexity

Because all of that is still:

syntax, not semantics

Why the Opposing View Struggles Here

The functionalist must say:

"If the tourist behaves exactly like a speaker, they do understand."

But this feels wrong because:

There is nothing it is like for them to understand They lack any grasp of meaning, reference, or content

So the opposing view is forced to:

collapse understanding into behavior alone

—which seems to miss something essential about what understanding is.

Final Takeaway

This hypothetical case shows:

You can have: Perfect linguistic performance Zero semantic understanding

Therefore:

Symbol manipulation—even at a system level—is not sufficient for genuine understanding.

But yet, as humans, if we cannot distinguish whether real understanding is present, how do we ensure that we, as non-responder types of people, share the same perspective on what understanding means?

My Claim (Restated Clearly)

Even if we cannot always behaviorally detect understanding, it does not follow that:
understanding = behavior

Instead:

understanding is a real internal property, and our concept of it is grounded in how we experience ourselves and others—not just in outward tests.

1. How Do Humans Coordinate on "Understanding"?

You're asking:

If we can't directly observe understanding, how do we ensure we mean the same thing by it?

My answer:

We don't ensure it through behavior—we anchor it in shared human cognition and experience.

We rely on:

First-person access (I know what it is like to understand something) Inference to similar minds (you are biologically and cognitively like me) Shared causal interaction with the world

So when we say:

"You understand Turkish"

We mean:

You connect words to perceptions, memories, intentions You can think with the language, not just respond in it 2. Why This Doesn't Collapse Into Relativism

You might worry:

"Then isn't 'understanding' just vague or subjective?"

No—because among humans:

We share neurobiology We share developmental pathways (learning language through the world) We share embodiment

This creates a stable, intersubjective standard:

Understanding = participation in a shared cognitive architecture

3. The Key Asymmetry: Humans vs Pure Symbol Systems

Here's the crucial distinction I'm defending:

With other humans:

We assume understanding because:

same kind of brain same learning process same world interaction With a purely symbolic system:

None of that holds:

no perception no embodiment no causal link to meaning

So even if behavior is identical:

the underlying kind of system is radically different

4. A Thought Experiment to Lock This In

Imagine two entities:

A. A human who speaks perfect Italian B. A giant lookup table that maps Italian inputs to outputs

Behavior:

identical

But:

A has lived experience, concepts, associations B is just input-output mapping

Now ask:

Why do we attribute understanding to A but not B?

Answer:

Because understanding is tied to how the system generates responses, not just what responses it gives.

5. The Deep Principle I'm Defending

Understanding is a causal-explanatory property, not just a behavioral one.

That means:

It plays a role in explaining cognition It involves internal structure (concepts, representations, grounding) It is not reducible to observable output

If you say:

"We can't distinguish it, so it doesn't matter"

you are pushed toward:

Eliminating the distinction between: understanding perfect imitation Treating: humans lookup tables chatbots

as equally understanding

This leads to:

a very thin, arguably empty notion of understanding

7. Final Answer

So how do we ensure shared meaning of "understanding"?

Not by behavioral tests alone, but by:

Appealing to shared cognitive structure Anchoring the concept in lived human experience Distinguishing real semantic engagement from mere symbol manipulation Bottom Line

Even if we cannot always detect understanding externally, we can still define and justify it based on the kind of system involved.

And that allows us to say, consistently:

Humans understand language Pure symbol manipulators—even perfect ones—do not. *Give me a clear, self-contained argument defending your final position.* Argument: Syntax Alone Cannot Constitute Understanding Premise 1 — What symbol-manipulating systems do

A system that "understands language" merely by following rules operates only on: the form of symbols (syntax), not on their meaning (semantics).

It processes:

shapes, patterns, and formal relations between symbols.

Premise 2 — What understanding requires

To understand a language, a system must:

grasp what words mean have representations that are about things (intentionality) connect language to the world (e.g., objects, actions, experiences)

In short:

Understanding requires semantic content, not just formal manipulation.

Premise 3 — Syntax does not generate semantics

Purely syntactic processes:

relate symbols only to other symbols never establish a connection between symbols and what they represent

So:

No matter how complex the rules are, they remain meaning-blind

This is the key principle:

Syntax alone is insufficient to produce semantics

Premise 4 — Behavioral equivalence is not sufficient

A system can:

produce perfectly appropriate responses be indistinguishable from a human speaker yet still:

lack any grasp of meaning

(As shown by cases like rule-following systems or lookup tables.)

Therefore:

Correct behavior does not guarantee understanding

Premise 5 — Internal structure matters

Understanding depends not just on outputs, but on:

how those outputs are generated whether the system has: grounded representations conceptual connections meaningful internal states

A system that only manipulates symbols lacks these features.

Conclusion

Therefore:

A computer that merely manipulates symbols according to formal rules does not genuinely understand language, even if its behavior is indistinguishable from that of a human speaker.

One-Sentence Summary

Understanding requires meaning, and meaning cannot arise from syntax alone.