How AI Vindicates Classical Theories of Meaning: Evidence from Large Language Models for the Semantics-Pragmatics Distinction

Abstract

This paper argues that the demonstrated capabilities of large language models (LLMs) provide surprising empirical support for classical theories of meaning, particularly the distinction between semantics and pragmatics and the reality of compositional literal meaning. While LLMs employ connectionist architectures rather than classical computational ones, their ability to systematically process novel sentences and distinguish between literal and contextual meaning suggests that key insights of classical semantic theory capture genuine features of linguistic understanding, even if the underlying mechanisms differ from those traditionally posited.

1. Introduction

Classical theories of linguistic meaning, as exemplified by philosophers and linguists in the tradition of Russell, Carnap, Chomsky, and Fodor, maintain several key distinctions and claims about the nature of linguistic understanding:

1. A fundamental distinction between semantics (literal meaning) and pragmatics (communicated meaning)

2. The reality of compositional literal meaning distinct from contextual interpretation

3. The systematic nature of linguistic understanding, particularly in handling novel sentences

These views have faced significant challenges from multiple directions. Speech-act theorists like Grice and Searle argue for the primacy of speaker intentions in determining meaning, while philosophers like Wittgenstein and proponents of discourse theory contend that meaning is inherently contextual. Both camps suggest that the classical view incorrectly posits unnecessary mental machinery and problematic Platonic entities. This paper argues that the demonstrated capabilities of large language models provide surprising empirical support for key aspects of the classical view, even while potentially requiring revisions to traditional assumptions about implementation.

2. Evidence from Novel Sentence Processing

Consider the sentence: "There exists a colorless green dream and it sleeps furiously." Despite its semantic anomalousness, LLMs can systematically:

1. Draw valid inferences from this premise (e.g., "there exists a dream," "something is both colorless and green")

2. Identify sentences that would entail it (e.g., "everything is a colorless green dream that sleeps furiously")

3. Process its logical and grammatical structure independently of its semantic coherence

This capability with novel sentences cannot be explained purely through statistical pattern matching of complete sentences, as the system has never encountered this exact formulation. Instead, it demonstrates understanding of:

- Component meanings

- Grammatical structure

- Compositional rules for combining these elements

This provides evidence for some form of compositional understanding, even if implemented differently than in classical computational models.

3. Convergent Paths to Compositional Understanding: Human and AI Language Acquisition

3.1 First Language Acquisition in Humans

Consider how a person X acquires their first language. Since X has no prior language to map new expressions onto, they must begin with direct encounters with language in use. The process appears to proceed as follows:

1. Initial Phase: X begins by grasping approximate communicated meanings in specific contexts. For instance, X learns that "that dog is going to bite you" sometimes indicates danger and calls for action.

2. Pattern Recognition: X notices systematic variations in:

- How the same sentence means different things in different contexts

- How similar sentences (e.g., "that dog is big") relate to each other

- How structurally analogous sentences (e.g., "this man is running towards you") function

3. Triangulation: Through exposure to an expanding set of sentences and their contextual meanings, X begins to:

- Refine understanding of how sentence meanings depend on their parts

- Notice how meaning differences correlate with component differences (e.g., how "John is a bad boy" differs from "Oliver is a bad boy" based on the name used)

- Develop increasing grasp of compositional structure

4. Emergence of Literal Meaning: Eventually, X develops the ability to:

- First assign a context-independent "default" meaning based on compositional structure

- Then modify this based on contextual factors to arrive at communicated meaning

3.2 Language Learning in AI Systems

AI systems like large language models learn language quite differently:

- Training exclusively on text data

- No direct access to situational contexts

- Learning purely from expression-to-expression patterns

- Massive parallel exposure rather than sequential development

3.3 Convergent Development of Compositional Understanding

Despite these profound differences in learning conditions and data sources, both humans and AI systems appear to develop similar capabilities:

1. Different but Related Data Sets:

- Humans learn from: (a) correlations between expressions and situations, and (b) correlations between expression-patterns and situations

- AI learns from: extensive patterns of expression-to-expression relationships

- Both are fundamentally statistical learning processes

2. Triangulation to Compositionality:

- Humans triangulate from contextual meanings to compositional understanding

- AI systems appear to derive similar compositional capabilities from purely linguistic patterns

- Both eventually develop ability to process novel sentences compositionally

3. Emergence of Literal Meaning:

- In both cases, grasp of context-independent literal meaning emerges from statistical pattern learning

- This suggests literal meaning is real but emergent, rather than primary

This convergence of capabilities from different learning paths provides strong evidence that:

- Compositional understanding is a robust feature of language processing

- Statistical learning can give rise to systematic understanding

- The distinction between literal and contextual meaning reflects something fundamental about linguistic meaning

4. Statistical Learning and Compositionality

4.1 Two Types of Statistical Data

Human language learners work with two distinct types of statistical patterns:

1. Expression-to-expression patterns (correlations between linguistic elements)

2. Expression-to-situation patterns (correlations between expressions and contexts of use)

While LLMs primarily access the first type of pattern, they appear able to derive much of what humans learn from situational context through rich analysis of purely linguistic patterns.

4.2 Emergence of Compositional Understanding

The fact that LLMs develop compositional capabilities through statistical learning suggests that compositionality need not require classical computational implementation. This indicates that the key insights of classical semantic theory about the nature of meaning can be separated from specific claims about implementation.

5. Addressing Critics

5.1 The Speech-Act Challenge

Perhaps the most significant challenge to classical views comes from speech-act theorists, particularly H.P. Grice and John Searle. Grice takes the radical position that sentence meaning is reducible to speaker meaning—what a sentence means is simply what speakers mean by it. Thus, if someone utters "one plus one equals two" intending to convey that political propaganda warps people's views, then according to Grice, that is what the sentence means in that context. Searle advances a similar but less extreme view, emphasizing the primacy of speaker intentions in determining meaning.

The speech-act perspective poses a particular challenge to the semantics-pragmatics distinction because it suggests there is no stable "literal meaning" independent of speaker intentions and context. However, the evidence from AI systems strongly counters this view:

1. LLMs can systematically process and understand novel sentences without any access to speaker intentions

2. They demonstrate stable compositional understanding across contexts

3. They can distinguish between literal content and contextual implications

This suggests that sentence meaning has a systematic structure independent of speaker intentions, even if those intentions play a crucial role in communication.

5.2 Wittgenstein and Discourse Theory

Wittgenstein and discourse theorists raise additional objections to classical views:

1. The positing of unnecessary "hidden depths" of mental computation

2. Commitment to problematic Platonic meanings

3. Inappropriate mathematization of language

However, the LLM example shows how systematic, compositional understanding can emerge without requiring:

- Classical computational architecture

- Platonic meanings beyond patterns of use

- Strict mathematical formalization

5.3 Synthesizing the Evidence

The capabilities of LLMs suggest a middle path that preserves key classical insights while acknowledging critics' valid points:

1. Against Grice and Searle:

- Sentence meaning cannot be reduced to speaker meaning, as evidenced by AI systems' ability to process meaning without access to speaker intentions

- However, their emphasis on the importance of communicative context remains valuable

2. Against Wittgenstein and discourse theorists:

- Systematic meaning structure exists, but need not require classical computational implementation

- Compositionality can emerge from pattern recognition rather than abstract rules

This evidence suggests that while meaning is deeply connected to use and context, it nonetheless has systematic structural features that cannot be reduced to speaker intentions or contextual factors alone.

6. Implications for Semantic Theory

The evidence from LLMs suggests several important conclusions:

1. Some form of the semantics-pragmatics distinction captures real features of linguistic understanding

2. Compositional literal meaning is real, even if implemented differently than traditionally assumed

3. Statistical learning can give rise to systematic understanding

4. Classical insights about the nature of meaning can be separated from classical claims about implementation

7. Conclusion

While AI systems may process language differently than humans and may learn through different mechanisms than those posited by classical theories, their demonstrated capabilities provide surprising support for key classical distinctions in semantic theory. This suggests that while traditional views may need revision, their core insights about the nature of linguistic meaning capture genuine features of language understanding.

These findings point toward a refined theory of meaning that preserves important classical distinctions while remaining open to new understanding of implementation mechanisms. This may help bridge the gap between classical semantic theory and more usage-based approaches to meaning.

References

[To be added based on specific works referenced]