

From Confusion to Clarity: The Power of Artificial Intelligence in Understanding Linguistic Ambiguities

Jumanah Abdulwahab Abusulaiman¹

¹Department of English Language, Umm Al-Qura University, Saudi Arabia

Correspondence: Jumanah Abdulwahab Abusulaiman, Department of English Language, Umm Al-Qura University, Saudi Arabia. E-mail: jaabusulaiman@uqu.edu.sa

Received: May 31, 2025

Accepted: August 25, 2025

Online Published: February 19, 2026

doi:10.5430/wjel.v16n3p273

URL: <https://doi.org/10.5430/wjel.v16n3p273>

Abstract

In this study I investigate how artificial intelligence (AI) identifies and analyzes two types of linguistic ambiguity: structural and semantic. I also seek to bridge the gap between theoretical linguistics and recent advancements in AI. I examine examples of linguistically ambiguous sentences using two AI applications, ChatGPT-4o and Google Gemini. Using a comparative analysis, I evaluate their comprehension and capabilities in resolving linguistic ambiguities. To enhance the evaluation, I use some prompts to elicit a second interpretation in cases where the applications fail to detect either structural or semantic ambiguities. The results reveal that ChatGPT-4o is more effective at identifying such ambiguities than is Google Gemini, which often struggles with ambiguous sentences. Semantic ambiguities, particularly those involving modals such as “must” and “ought to,” or scope issues with quantifiers such as “every,” pose greater challenges for both applications than do structural ambiguities. Despite these challenges, both ChatGPT-4o and Google Gemini demonstrate useful disambiguation strategies, including referencing the context of the utterance and paraphrasing ambiguous sentences to clarify their meaning. The study highlights the significant potential of AI in imitating human capabilities related to sentence comprehension and the resolution of linguistic ambiguities.

Keywords: structural ambiguity, semantic ambiguity, artificial intelligence, ChatGPT-4o, Google Gemini

1. Introduction

Linguistic ambiguities inevitably arise during human communication and are a main challenge to comprehension. Ambiguity occurs when a person interprets a single sentence in two ways (paraphrases) caused by false reasoning. However, this linguistic phenomenon also reveals an important feature of the human mind for entertaining these ambiguities. Further, the human mind can disambiguate them relative to an appropriate context.

In this study I do not examine linguistic ambiguities from a traditional linguistic perspective because researchers analyze these types of ambiguities in the linguistic literature. Instead, I integrate them with recent advancements in artificial intelligence (AI) and investigate how AI interprets and interacts with such ambiguities. These two fields currently operate largely independently, with central themes in theoretical linguistics rarely integrated into AI research. As recent developments in AI emerge, though, many second language learners who encounter difficulties in understanding linguistic ambiguities often turn to AI applications for assistance. For example, ChatGPT-4o has proven to be a valuable resource for interpreting linguistically ambiguous sentences. As discussed later, in Section 4.1, although ChatGPT-4o is effective at identifying ambiguities, it occasionally struggles to comprehend certain sentences fully.

Since the emergence of AI in 2022, no researchers have examined how it comprehends structural and semantic ambiguities; therefore, whether AI can both identify ambiguity and provide alternative interpretations of ambiguous sentences remains unclear. To the best of my knowledge, Ortega-Martín, Sierra, Ardoiz, Álvarez, Armenteros, & Alonso (2023) conducted the first study addressing this phenomenon. This highlights a gap in the literature of linguistics because researchers have not reexamined linguistic ambiguity in light of recent AI developments. Moreover, the phenomenon has not received adequate attention from linguists and philosophers reexamining this linguistic phenomenon to accommodate recent advancements in technology. In this study I seek to establish a new perspective on linguistic ambiguity from a natural language processing approach, linking it to a long-standing phenomenon in logic and linguistics.

I intend the following questions to open avenues for future researchers to explore structural and semantic ambiguities relative to AI.

1. Can AI detect structural and semantic ambiguities in sentences?
2. Is AI capable of generating a second interpretation of ambiguous sentences? How does it interpret such sentences?
3. Can AI provide examples of structurally and semantically ambiguous sentences?
4. Can AI resolve structural and semantical ambiguities in the same manner as humans can?

I organize the remainder of this paper as follows. In Section 2, I present the theoretical framework and show how philosophers and

linguists explain structural ambiguities and their approaches to interpreting semantic ambiguities. This section reflects how humans comprehend and interpret structural and semantic ambiguity. In Section 3, I introduce the method I employ in data collection, research questions, and the use of prompts. In Section 4, I examine and analyze how AI interprets the same examples of structural and semantic ambiguities previously discussed in Section 2. In Section 5, I discuss key findings of the comparison between the two AI applications explored here. In Section 6, I conclude the discussion and note the study's limitations.

2. Theoretical Framework

In the linguistic ambiguities literature, philosophers and linguists agree on the classification of linguistic ambiguities into lexical, phonological, semantic, and structural ambiguities (Gillon, 2019; Hurford, Heasley, & Smith, 2007; L öbner, 2014). Regardless of the type of linguistic ambiguity, undeniably, such ambiguity represents a natural phenomenon of language arising from human communication. A sentence remains unclear unless the listener understands the pragmatic elements surrounding a situation. In this study, I focus primarily on only two types of linguistic ambiguity: structural and semantic. Let us begin the discussion with the following example, which represents structural ambiguity:

(1) I watched the man with the telescope.

In example (1), the structural ambiguity is triggered by the phrase “with the telescope,” which has two possible interpretations. One is that the man referred to in the sentence is holding the telescope. The other is that the person referred to as “I” is watching the man through a telescope. In other words, the different scope effects of the prepositional phrase (PP) “with the telescope” causes the structural ambiguity. In the first case, the PP modifies the determiner phrase (DP) “the man,” who is the one holding the telescope, whereas in the second case, the PP is attached to the main verb “watched” and describes the event of watching the man.

Because this type of ambiguity frequently occurs during social interactions among speakers, it presents challenges for nonnative speakers of English. Therefore, it is a linguistic and language acquisition problem for many English learners. One technical method for resolving the structural ambiguity in example (1) is to use square brackets to clarify phrase structure inside the sentence.

(1a) [I watched [the man [with the telescope]]]

(1b) [I watched [the man] [with the telescope]]

Semantic ambiguity, in contrast, can be caused by the presence of a pronominal, such as “everyone,” which causes difficulty for creating a reference to an entity in a sentence, as presented in example (2):

(2) Everyone loves his mother.

In example (2), the pronoun “everyone” could refer to John’s mother as “his mother,” with every person loving her. Another interpretation arises when the possessive pronoun “his” is co-indexed with “everyone” to yield different mothers.

Allan (2001) proposed a feasible way to resolve the linguistic ambiguity in a sentence by associating the sentence with the time and world pair of an eventuality (real or imaginary). In semantics, world and time refer to “the context of utterance” (Hurford et al., 2007, p. 71). Without considering the context of the utterance associated with the sentence, we cannot determine an appropriate interpretation of any ambiguous sentence.

In addition, when a speaker produces a sentence, the context includes their attitudes, emotions, and beliefs. We cannot ignore the effect of these factors on both the sentence semantics and sentence interpretation. This is a distinctive feature of human communication, one not available to AI unless experts and specialized machines test it.

Examples (1) and (2) demonstrate that disambiguating structural and semantic ambiguities is a distinctive property of human cognitive processes. In this study I investigate resolving these ambiguities through the use of two AI applications, ChatGPT and Google Gemini, rather than relying on linguistic literature.

2.1 Structural Ambiguities

This section offers a brief overview of structural ambiguity, known in logic as “amphiboly.” One cannot judge a sentence true or false until they have evaluated it within the context of the utterance given the appropriate world and time. The following are additional examples of structural ambiguity:

(3) John and Mary are married.

Example (3) is ambiguous due to the conjunction “and,” which yields two interpretations. The first suggests that John and Mary are married to each other, and the second suggests that John and Mary are married to different spouses. Only pragmatic elements can help us understand the appropriate interpretation.

Here are additional examples of sentences with structural ambiguities:

(4) I read the book on the balcony.

In example (4), the speaker refers to a reading eventuality occurring on a balcony. However, the prepositional phrase “on the balcony” either can modify the place of the verb phrase “read” or can be attached to the noun phrase “the book” and modify the place of the book in the sentence. The famous comedian Groucho Marx spoke another example of ambiguity:

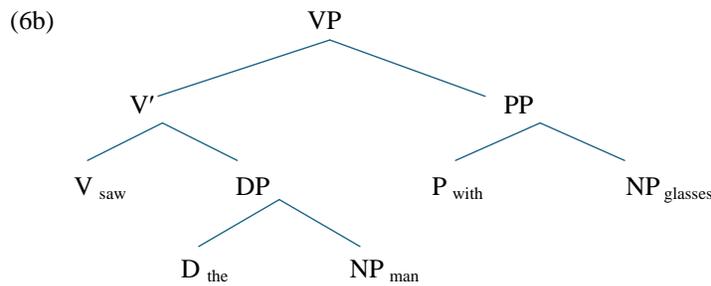
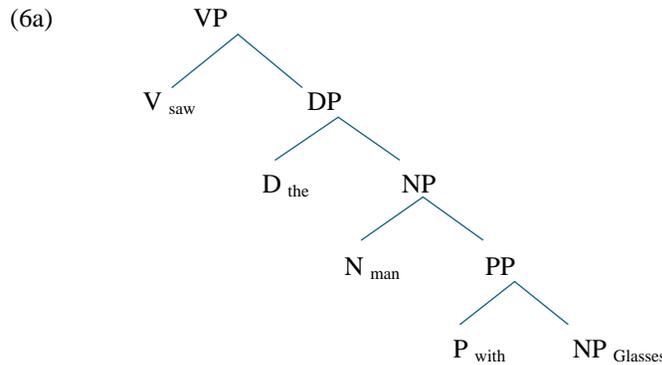
(5) One morning, I shot an elephant in my pajamas.

Example (5) is structurally ambiguous because of the noun phrase “in my pajamas.” The first interpretation refers to a person wearing pajamas shooting an elephant, and the second interpretation suggests that the elephant was wearing the shooter’s pajamas.

Following is another instance of a structurally ambiguous sentence in (6), as syntactically presented in (6a) and (6b) by van Gelderen (2010, p. 41):

(6) She saw the man with glasses.

Example (6) is structurally ambiguous because of the PP “with glasses,” which can be interpreted in two ways. One interpretation is that the PP attaches to the noun phrase (NP) “the man,” as illustrated in (6a). The other interpretation is to anchor the PP “with glasses” to the verb phrase (VP) “saw,” as shown in (6b). Here is a syntactic representation that clarifies this structural ambiguity:



The preceding discussion illustrates how a human being interprets structural ambiguities, and I examine and analyze how AI detects such ambiguities in Section 4. Meanwhile, the following section provides additional examples of semantic ambiguities and the manner in which a human is capable of interpreting them. As stated earlier, world, time, and attitude can all contribute to sentence comprehension. We judge the sentence as felicitous or infelicitous under the appropriate circumstances.

2.2 Semantic Ambiguities

In this section, I discuss the way linguists describe and treat semantic ambiguity. Semantic ambiguity is usually triggered by using a referring expression such as the pronominal “he,” or a quantificational expression such as “every,” in a sentence or utterance. Each interpretation of a sentence represents a different truth condition, which is determined as either true or false by fixing the interpretation of the referring expression (Poesio, 1996). The context of the utterance can help determine the reference when producing a sentence. Lyons (1977, p. 661) gave an illustrative example of semantic ambiguity with a pronoun:

(7) John looked up when he came in.

It is unclear whether the pronoun “he” refers to John as an antecedent in the sentence or someone other than John. The pronoun “he” might be interpreted as dietic or anaphoric. If interpreted as a direct marker, it refers to someone else present in the context of utterance by a head nod or a hand gesture. If “he” is treated as anaphoric, then it is co-indexed with John or an antecedent in the preceding sentence. The example presented in example (7) can be interpreted only by pragmatic elements and the context of the utterance but not through the structure of the sentence.

Following is another example that represents semantic ambiguity triggered by the pronoun “every”:

(8) Every cheerleader had shared a milkshake with Jones.

In example (8), one interpretation is that there are different milkshake containers shared by “every cheerleader” with Jones. The other interpretation refers to the same milkshake container that is shared by “every cheerleader” with Jones.

The observation reported in example (8) represents a well-known semantic phenomenon known as “scopal ambiguity” (Kearns, 2011, p. 50). Let us apply this scope method to the following example:

(9) The company sent one representative to every meeting.

In example (9), imagine a scenario with one company, two representatives (R_1 and R_2), and three meetings (M_1 , M_2 , and M_3). The company sent R_1 to M_1 and M_2 and R_2 , to M_3 . This interpretation suggests that different company representatives attended every meeting, whereas the other interpretation proposes that only one representative (R_1) attended all the meetings.

(10) The chicken is ready to eat.

This is semantically ambiguous because the speaker may be referring to a chicken ready to be fed on his farm or that the chicken is cooked and ready to be eaten.

Having examined how human beings interpret examples of structural and semantic ambiguities, I now explore in the following section how the two AI applications, ChatGPT-4o and Google Gemini, deal with the same examples.

3. Methodology

3.1 Linguistic Ambiguities and Artificial Intelligence

To investigate how AI recognizes and processes ambiguous sentences in English, I conducted a comparative analysis on two types of AI applications, ChatGPT-4o and Google Gemini. It is important to highlight that no participants were involved, and I expected no numerical results from input–output processes. The data used in this paper are from linguistically ambiguous sentences identified in the existing literature. I cited some examples from linguistic sources for reference, while selecting others from the *Stanford Encyclopedia of Philosophy* (Sennet, 2023). To analyze linguistic ambiguities, I consolidated two fields—AI and theoretical linguistics—in this study. It is important to note that I do not investigate computational techniques in the field of AI; rather, I shed light on a well-known linguistic phenomenon in relation to AI.

To obtain additional illustration from ChatGPT-4o and Google Gemini, I applied a few prompts when the applications did not initially recognize sentences as unambiguous. The aim of prompts is to improve and narrow down the output that AI provides (Ortega-Martín et al., 2023). It is important to note that I reproduced the outputs presented in this study exactly as ChatGPT-4o and Google Gemini generated them.

Following a reviewer’s recommendation in a previous submission, it is important to document the date upon which I exploited this study’s data. I started investigating this phenomenon on October 2, 2024. Appendix A presents a complete list of tested sentences in these two applications.

In the following Section 4, I compare these two AI applications to determine how each can detect and address structural and semantic ambiguities. In particular, I investigate the following questions in the comparative analysis:

1. Is this sentence “ ” ambiguous?

This question does not concern the computational processes for structural and semantic ambiguities but simply tests whether AI can recognize such ambiguities in the same manner as humans can.

After receiving the output from AI for the first question, my first attempt is to prompt AI to evaluate its comprehension level; the second question represents this attempt:

2. Is there a second interpretation of this sentence “ ”?

I applied my second attempt to see the potential of AI in providing linguistically ambiguous sentences by itself, and the prompt is as follows:

3. Can you provide examples of structurally and semantically ambiguous sentences?

Findings are based on the aforementioned questions and their outputs that AI applications provided:

4. Can AI resolve structural and semantic ambiguities in the same manner as humans can?

4. Results and Analysis

In this section I report key findings obtained from testing structural and semantic ambiguous sentences via two applications of AI. I present all responses offered by ChatGPT-4o and Google Gemini here, without modifications, to maintain clarity. Section 4.1 provides responses from ChatGPT-4o, whereas Section 4.2 shows responses from Google Gemini.

4.1 ChatGPT-4o

I provided a series of sentences to ChatGPT-4o, asking whether they were ambiguous. When ChatGPT-4o considered a sentence unambiguous, I used additional prompts to test its comprehension, such as “Is it possible to have a second interpretation of this sentence?” I have reproduced the outputs exactly as ChatGPT-4o provided, with no modifications. Let us consider the following examples:

Input (1): Is the sentence “The old men and women left the room” ambiguous?

ChatGPT-4o output (1): Yes, the sentence “The old men and women left the room” can be considered ambiguous because it is unclear

whether “old” modifies both “men” and “women” (implying both are old) or just “men” (implying only the men are old while the women may be of any age). Therefore, without further context, the sentence does not clearly specify the ages of the women.

This output demonstrates ChatGPT-4o’s ability to analyze ambiguity in sentences. Its recognition that context is necessary to resolve ambiguity is particularly interesting. As discussed earlier, pragmatic elements are essential in disambiguating sentences. Following are additional examples to support this recognition by ChatGPT-4o:

Input (2): Is the sentence “The man eating the fish is sick” ambiguous?

ChatGPT-4o output (2): Yes, the sentence “The man eating the fish is sick” is ambiguous because it capable of two interpretations:

- (a) The man, who is currently eating fish, is ill.
- (b) The man has eaten fish that is spoiled or contaminated, which has caused his illness.

The ambiguity here arises from whether “sick” refers to the man’s current health or to the condition of the fish he is eating; context would clarify the intended meaning.

ChatGPT-4o not only succeeds in disambiguating sentences with syntactic ambiguity but also insists on the role of context in resolving the ambiguity. This is a remarkable property of ChatGPT-4o, to understand the sentence and analyze it in a manner similar to humans.

In the case of semantic ambiguity, ChatGPT-4o determined the sentence to be unambiguous in its output. In this case, I applied the following prompts:

Input (3): Is the sentence “You must eat a piece of cake” ambiguous?

ChatGPT-4o output (3): The sentence “You must eat a piece of cake” is not inherently ambiguous because it conveys a clear directive or obligation for the person addressed to eat a piece of cake. However, the word “must” can imply different degrees of necessity, which could introduce a slight nuance:

1. Obligation: It could imply a strong obligation, as in “It is required that you eat a piece of cake.”
2. Recommendation: It could suggest a strong recommendation or encouragement, implying that the speaker believes eating the cake would be a good idea. Overall, while there is room for interpretation regarding the strength of the requirement, the sentence’s main message is clear.

Here, ChatGPT-4o fails to identify the deeper semantic ambiguity that the modal “must” triggers. However, it admits that the modal “must” has different degrees of ambiguity: obligation and recommendation.

In contrast, in the literature of modality, the modal “must” has multiple interpretations in English, according to Kratzer (1977). The reader can determine the correct interpretation of the modal depending on the relevant context. This may refer to an obligation or a preference.

Another example of a semantic ambiguity is the following:

Input (4): Is the sentence “The company sent one representative to every meeting” ambiguous?

ChatGPT-4o output (4): The sentence “The company sent one representative to every meeting” is not inherently ambiguous, but it can lead to slightly different interpretations depending on the context:

1. One representative per meeting: This interpretation suggests that for each meeting, the company sent exactly one representative, implying that the representatives could be different individuals for different meetings.
2. The same representative for every meeting: It could also imply that the same representative attended all the meetings, indicating consistency in representation.

Although the first interpretation is a more common understanding, context would help clarify whether the representative was different for each meeting or if the same person attended all the meetings. Thus, despite the potential for varying interpretations, the sentence is generally clear regarding the action being described.

ChatGPT-4o’s strong specification for detecting and comprehending structurally ambiguous sentences, such as “The old men and women left the room” and “The man eating the fish is sick,” is quite remarkable. It efficiently resolves the ambiguity and interacts with prompts. Similarly, its recognition of the context required to disambiguate a sentence is also outstanding. Across all types of ambiguous sentences, the term “context” was present in all its outputs. However, semantic ambiguities seem to be challenging for ChatGPT-4o, as noted in the aforementioned examples. With semantic ambiguities, all its outputs assert the need for context as a pragmatic tool to determine the correct interpretation. This essential specification of ChatGPT-4o demonstrates its high potential as an AI application that imitates human beings in language processing.

Finally, I tested ChatGPT-4o to see if it can provide both structurally and semantically ambiguous sentences. Here are some of the examples provided by ChatGPT-4o:

Input (5): Provide some examples of structurally ambiguous sentences.

ChatGPT-4o output (5):

- (11) “Flying planes can be dangerous.”

Interpretation 1: The act of flying planes is dangerous.

Interpretation 2: Planes that are flying can be dangerous.

(12) “They chased the boy with a bike.”

Interpretation 1: They used a bike to chase the boy.

Interpretation 2: They chased a boy who had a bike.

Semantically ambiguous sentences:

(13) “They are cooking apples.”

Interpretation 1: The apples are intended for cooking.

Interpretation 2: Someone is cooking apples.

(14) “She didn’t eat all of the cookies.”

Interpretation 1: She ate none of the cookies. (Negation applies to all.)

Interpretation 2: She ate some but not all the cookies. (Negation applies to some.)

Interestingly, ChatGPT-4o could provide sentences with the modal “must” and recognize the ambiguity that arises with it, as shown in example (15):

(15) “Nobody must fail the exam.”

Interpretation 1: No one is allowed to fail. (Negation applies to the necessity of failing.)

Interpretation 2: It is not necessary that anyone fail (implying that some may fail, but it is not required). (Negation applies to the necessity.)

Further, ChatGPT-4o offers helpful techniques for resolving ambiguities, such as using context, rephrasing sentences, providing additional information, clarifying pronouns or referential terms, disambiguating scope, and using specific vocabulary.

In summary, ChatGPT-4o successfully detects both structural and semantic ambiguities as well as comprehends and analyzes such sentences with clarity, in a manner similar to human beings. In addition, ChatGPT-4o asks questions about whether the answers are clear enough or more clarification is necessary. Finally, ChatGPT-4o answers prompts, such as providing examples of ambiguous sentences and ways of resolving those ambiguities.

4.2 Google Gemini

In this section I present how Google Gemini detects and interprets structural and semantic ambiguities using the same examples tested with ChatGPT-4o. Let us start with examples where Google Gemini succeeds in resolving the ambiguities and then proceed to the ones where it fails.

Input (6): Is the sentence “The old men and women left the room” ambiguous?

Google Gemini output (6): Yes, the sentence is ambiguous. Here are the two possible interpretations: All the people who left the room were both old and men and women. This is a rather unlikely scenario, but it is a possible interpretation. The group consisted of old men and women of any age: This is the more likely interpretation. To avoid ambiguity, you could rephrase the sentence as “The elderly men and women left the room” (clearly indicating age for both genders). “The old men and the women left the room” (clearly indicating two separate groups).

Input (7): Is the sentence “The man eating the fish is sick” ambiguous?

Google Gemini output (7): Yes, the sentence “The man eating the fish is sick” is ambiguous. It is capable of two interpretations: The man who is currently eating the fish is sick. The fish that the man is eating is sick. To clarify the meaning, you could say, “The man who is eating the fish is ill.” “The fish the man is eating is not fresh.”

If we focus on the answer to input (6), Google Gemini offers a helpful technique for disambiguating a sentence other than the context of an utterance. It suggests paraphrasing the sentence for clarity. ChatGPT-4o does not offer this technique. ChatGPT-4o asserts the importance of the context of the utterance to disambiguate the sentence. This is an interesting difference between the two programs.

However, Google Gemini failed to detect ambiguities in a few sentences (Appendix B provides the full list).

(16) They are hunting dogs.

(17) I read the book on the balcony.

(18) I saw a cat with one eye.

(19) She knew a little Greek.

Let us examine how Google Gemini analyzes a few of these. For example (16), Google Gemini could not comprehend the ambiguity behind the noun phrase “hunting dogs” when asked the following question:

Input (8): Is the sentence “They are hunting dogs” ambiguous?

Google Gemini output (8): No, the sentence “They are hunting dogs” is not ambiguous. It clearly states that the subjects (referred to as “they”) are dogs used for hunting.

Based on the output in (8), I prompted Google Gemini to examine its comprehension of the noun phrase “hunting dogs.”

Input (9): What is meant by “hunting dogs” in “They are hunting dogs”?

Google Gemini output (9): Hunting dogs are dogs specifically trained and bred to assist humans in hunting. They have various roles depending on the type of hunt, such as the following:

- a. Retrieving: Bringing back game that has been shot
- b. Pointing: Indicating the location of prey
- c. Flushing: Driving game out of hiding
- d. Hounding: Pursuing and cornering prey
- e. Terrier: Digging out prey from burrows

Examples of hunting dog breeds include Labrador Retrievers, German Shorthaired Pointers, English Setters, Beagles, and Dachshunds.

Would you like to explore different scenarios or phrases related to hunting dogs?

I then asked a few prompts to lead to the second interpretation of example (16):

Input (10): Is it possible to refer to dogs that are actively hunting?

Google Gemini output (10): You could use phrases such as the following:

- a. Hunting dogs (general term for dogs used for hunting)
- b. Hounds on the hunt (specifically for dogs that chase and track prey)
- c. Dogs in pursuit (more general term for dogs actively chasing something)

Input (11): What are other possible interpretations of this sentence?

Google Gemini output (11): The sentence “They are hunting dogs” has a single, straightforward interpretation: The subjects referred to as “they” are dogs specifically bred or trained for hunting. There is no ambiguity in this sentence. It clearly states that the subjects discussed are dogs with a particular purpose: hunting.

At this level, Google Gemini could not identify the ambiguity of “hunting dogs” understood by humans, even after it asks, “Would you like to explore other sentences or topics?”

Surprisingly, Google Gemini decided that the sentence “I read the book on the balcony” was not ambiguous, whereas ChatGPT-4o immediately decided it was ambiguous. Here is the output received from Google Gemini for this sentence:

Input (12): Is the sentence “I read the book on the balcony” ambiguous?

Google Gemini output (12): No, the sentence “I read the book on the balcony” is not ambiguous. It clearly implies that the act of reading the book took place on the balcony.

Input (13): What is “on the balcony” in the sentence “I read the book on the balcony”?

Google Gemini output (13): The sentence “I read the book on the balcony” does not explicitly state what else is on the balcony. It only indicates that the act of reading the book occurred on the balcony. There could be many other things on the balcony, such as furniture (chairs, tables), plants, and other objects (e.g., decorations, toys). The sentence simply focuses on the location of the reading activity.

Google Gemini output (14): Would you like to try another sentence?

I further tested Google Gemini by asking whether there is a possible second interpretation of the sentence:

Input (15): Does this sentence have a second interpretation?

Google Gemini output (15): No, the sentence “I read the book on the balcony” does not have a second interpretation in the typical sense. It clearly and directly conveys the action of reading a book in a specific location. There is no room for ambiguity in this particular case.

Google Gemini output (16): Would you like to try another sentence?

Now, it becomes obvious that Google Gemini cannot capture the structural ambiguity in “I read the book on the balcony.” It fails to recognize the possibility that the book happens to be on the balcony and the speaker read it.

When it comes to semantic ambiguities, Google Gemini could not identify the ambiguity in examples with the modal “ought to”:

Input (17): Is the sentence “John ought to be at home by now” ambiguous?

Google Gemini output (17): No, the sentence “John ought to be at home by now” is not ambiguous. The phrase “ought to be at home by now” expresses a strong expectation or belief that John is currently at home. There is no other possible interpretation of this sentence.

This output eliminates the possibility for a second interpretation. I also examined whether there is a second interpretation of the modal “ought to” in input (17), as von Stechow and Iatridou (2008) described. They distinguished between “ought to” and “have” (like “must”) in examples such as “everybody ought to wash their hands” and “employees must” (p. 116).

Input (18): Is it possible to have a weak necessity interpretation of the modal “ought to” in this sentence?

Google Gemini output (18): Yes, it is possible to have a weak necessity interpretation of the modal “ought to” in the sentence “John ought to be at home by now.” Although a sentence primarily conveys a strong expectation or belief, it can also be interpreted as a weak obligation or suggestion. In this case, the sentence implies that it would be advisable or preferable for John to be at home; however, this is not strictly necessary. This interpretation is often used when there are no serious consequences of not following the suggestions. For example, if John is expected to be home for dinner but there is no urgent need for him to be there, the sentence could be interpreted as a weaker obligation.

Given Google Gemini’s output in (17), surprisingly, it offers another interpretation of the modal “ought to” via a prompt. ChatGPT-4o easily detected the same sentences and provided the following answers:

Yes, the sentence “John ought to be at home by now” is ambiguous due to a likely typographical error. It seems to be missing a verb, possibly “should” or “ought,” which could lead to confusion. If we correct it to “John should be at home by now” or “John ought to be at home by now,” it becomes clearer. However, the ambiguity primarily arises from the implication of what “by now” means. The phrase could suggest (1) John is expected to be at home based on the time elapsed or (2) circumstances may have changed that could influence whether John is at home. While the original sentence lacks clarity because of the missing word, the intended meaning may still present some ambiguity on the context and the specific situation being discussed.

Here is another example with semantic ambiguity that Google Gemini failed to detect:

Input (19): Is the sentence “You must eat a piece of cake” ambiguous?

Google Gemini output (19): No, the sentence “You must eat a piece of cake” is not ambiguous. The modal verb “must” clearly expresses a strong obligation or necessity. No other reasonable interpretations exist in this context.

Interestingly, Google Gemini described what is meant by structural ambiguities when it was prompted. It also provided a few sentences with structural ambiguities, such as *I saw the man with a telescope*, *the student was attacked by the bear in the forest*, and *visiting relatives can be boring*. These examples are similar to the ones used as inputs by the researcher. They all have issues with prepositional phrase attachments.

With respect to semantic ambiguities, Google Gemini offered a few examples with referential (pronouns), scope (quantifier), and lexical ambiguities like *I went to the bank*, *everybody isn’t here*, *Bob kicked Tom*, and *he broke his leg*. This AI application also explained the semantic reasons for causing semantic ambiguities in each sentence. However, it did not provide any example with a modal ambiguity.

5. Discussion

In this section, I discuss principal results of the comparison between ChatGPT-4o and Google Gemini relative to structural and semantic ambiguities.

For ChatGPT-4o, structural ambiguities were easy to detect. This AI application could directly offer two interpretations of structurally ambiguous sentences, specially sentences that include adjectival phrases, as in *old men and women*, *flying planes*, *with a bike*, or *cooking apples*. In addition, ChatGPT-4o offered a strategy to resolve the ambiguities in these sentences, such as further contextual elements. The offered strategy by ChatGPT-4o is to specify the age of men and women, so the structural ambiguity can be resolved. This distinctive property of ChatGPT-4o of resolving structural ambiguity and offering a helpful way of resolving the ambiguity truly creates a trust relationship with its users.

As regards semantic ambiguities, I tested sentences with the English modal “must” and a quantifier “every.”

With the modal “must,” ChatGPT-4o unfortunately failed to determine semantic ambiguities in sentences with the English modal “must” in *you must eat a piece of cake*. It offered only one response or interpretation, which is associated with strong obligation. It did not recognize another possible interpretation of the modal “must” even after using prompts by the researcher.

To return to the literature of formal semantics, Kratzer (1977) claimed that the modal “must” has multiple interpretations ranging from circumstantial to epistemic interpretations. The modal “must” obtains its interpretation on the basis of the conversational background that is compatible with the modal to yield “epistemic necessity reading,” “a root deontic reading,” or “a preferential reading.” The multiple readings of the modal “must” are clearly represented by Hacquard (2016).

- | | |
|--|---------------------|
| (20) a. John is not in his office. He must be home. | <i>Epistemic</i> |
| b. John parked illegally. He must pay a fine. | <i>Deontic</i> |
| c. John wants to get a PhD. He must write a thesis. | <i>Teleological</i> |

(Hacquard, 2016, p. 45)

Knowing the conversational background is a crucial element to determine the reading of the modal “must”. I believe that understanding the conversational background of this modal or any other modal is a human privilege that AI applications poorly exhibit with modals, such as “must”. Remember that ChatGPT-4o cannot recognize this fact about English modals. Only humans can figure out the pragmatical elements surrounding an event or a modal. A speaker and a listener should have mutual understanding if they share the same conversational background. Accordingly, this shared understanding leads to a successful communication. This contextual requirement is an essential requirement in any human conversations and a reminiscent of Grice’s (1969) discussion about successful communication between speakers.

With the quantifier “every,” ChatGPT-4o did not detect the semantic ambiguity in a sentence like *the company sent one representative to every meeting*. Its response was *it is not inherently ambiguous*. Even though it rejected the ambiguity in this sentence, it could explain two interpretations of the sentence caused by the presence of the quantifier “every”

I find that resolving semantic ambiguity is considered a shortcoming of ChatGPT-4o. This finding offers a useful direction for the developers of ChatGPT-4o to follow up on in the future. ChatGPT-4o or other types of AI application should be developed to accommodate contextual elements surrounding examples with modals and quantifiers.

For Google Gemini, it easily identified structural ambiguities in *the old men and women*. Interestingly, it offered two paraphrases to disambiguate the sentence. For the first paraphrase to indicate age for both genders, Google Gemini recommended this paraphrase: *The elderly men and women . . .* I find this paraphrase is still ambiguous itself because it could refer to both elderly men and women. For the second paraphrase to refer to separate groups, Google Gemini suggested this paraphrase: *The old men and the women . . .*

Even though Google Gemini offered two paraphrases for resolving this structural ambiguity, the first paraphrase is ambiguous itself, as already discussed. This finding invites a recommendation that a human should reexamine this sentence and should not completely depend on the output provided by Google Gemini or any other AI application.

Another interesting observation of Google Gemini is offering another paraphrase for this sentence: *The man eating the fish is sick*. Again, offering a paraphrase is a property that Google Gemini is capable of. ChatGPT-4o, conversely, insisted only on knowing the context of the utterance, without providing any paraphrase for structural ambiguity.

Unfortunately, Google Gemini could not detect a few structurally ambiguous sentences, such as *I read the book on the balcony*. Therefore, I needed prompts at this level to trigger more outputs from Google Gemini. The only reading that Google Gemini provided was relevant to the location of reading the book.

Similar to ChatGPT-4o, Google Gemini failed to identify semantically ambiguous sentences with English modals like “ought to” as in *John ought to be at home by now*. Google Gemini only recognized the strong obligation reading of “ought to”. After applying several prompts, Google Gemini acknowledged “a suggestion reading” of “ought to”.

In natural human communication, English modals, such as “ought to,” are claimed to have multiple readings ranging from epistemic, goal-oriented, and deontic readings (von Fintel & Iatridou, 2008).

- (21) It’s 3pm. Given what I know about Morris’s habits, he **ought to** be in his office. Epistemic
(von Fintel & Iatridou, 2008, p. 126)
- (22) To go to Ashfield, you **ought to** take Route 2. Goal-oriented
- (23) You **ought to** do the dishes. Deontic
(von Fintel & Iatridou, 2008, p. 127)

Given the preceding discussion, both AI applications—ChatGPT-4o and Google Gemini—share weakness in interpreting semantically ambiguous sentences. This observation reveals a distinctive fact of a human mind and its ability to recognize and understand such ambiguous sentences with modal verbs. As claimed by Kratzer (1977), von Fintel and Iatridou (2008), and many others, modal verbs acquire multiple readings depending on the conversational background or even the aspectual morphology of the clause following them (see Condoravdi [2001], Bhatt [2006], and Hacqurad [2006], among many other semanticists).

Table 1 summarizes key findings of the comparison between ChatGPT-4o and Google Gemini relative to structural and semantic ambiguities.

Table 1. A comparison between ChatGPT-4o and Google Gemini

	Structural ambiguity	Semantic ambiguity	Offering ambiguous sentences
ChatGPT-4o	Insisted on the role of further context	Failed to identify semantically ambiguous sentences	√
Google Gemini	Offered two paraphrases of a structural ambiguous sentence		√

The discussion herein leads us to acknowledge the importance of AI applications in life. At the same time, humans have a distinctive property in understanding language as an innate ability of using a natural language. After discussing the challenge of detecting semantic ambiguities with two instances of AI applications, users should not solely depend on the use of AI. They should refer to other sources in the linguistic literature or ask native speakers to resolve linguistically ambiguous sentences.

In addition, the weak manifestation that ChatGPT-4o and Google Gemini reveal in detecting semantic ambiguity should be revisited by AI application developers. Such applications might give misleading outputs for users specially those whose native language is not English, or they are not linguists.

To sum, Both AI applications responded successfully when they were prompted to give syntactic and semantic ambiguous sentences. They correctly interpreted the ambiguity exhibited in each example. Unfortunately, ChatGPT-4o and Google Gemini's outputs examples were very limited.

6. Conclusion

Recent advancements in AI call for a reexamination of linguistic ambiguities. In this study I investigated how AI detects both structural and semantic ambiguities and contributes to theoretical linguistics by offering a new perspective on detecting structural and semantic ambiguities by establishing a dialogue with two common applications of the AI field, ChatGPT-4o and Google Gemini.

I found that AI has greater potential for detecting structural than semantic ambiguities. Semantic ambiguities seem to be challenging to identify, especially those sentences with modals, such as "must" and "ought to," and with scope problems, such as "every."

Having examined two AI programs, I found ChatGPT-4o outperformed Google Gemini in most ambiguous sentences, even those with modals. It also offered helpful techniques to disambiguate sentences and proposed questions for clarity, such as "Is this answer enough for you?"

Google Gemini, by contrast, often failed to provide second interpretations of a group of sentences with both structural and semantic ambiguities. Even after I used various prompts to trigger second interpretations of those sentences, it was still unable to detect their ambiguities.

Given the aforementioned circumstances, ChatGPT-4o has proven to be a helpful resource for users, especially nonnative speakers with intermediate levels of English as a second language, as a research tool. I note, however, that users should not completely depend on AI tools, including ChatGPT-4o, to disambiguate linguistic ambiguities. Traditional research methods, such as books and Google Scholar, remain valuable resources for addressing linguistic ambiguities. However, further investigation into these ambiguities within the realm of AI is necessary. In future researchers should focus on enhancing the analysis and detection of linguistic ambiguities. Many second language learners rely on ChatGPT-4o to navigate these challenges. As learners work on developing their second language skills, it is essential to improve ChatGPT-4o and similar AI applications to better accommodate the natural language used in everyday communication.

Acknowledgement

I would like to express my sincere appreciation to the anonymous reviewers for their time, effort, and insightful feedback. Their constructive comments and valuable suggestions have significantly improved the clarity and quality of this paper. I also extend my thanks to Hanadi Azhari for her constructive insight and to the members of Translation and Arabization Center at Um Al-Qura University for the fruitful discussions on the role of AI in translation.

Authors' contributions

Not applicable

Funding

Not applicable

Competing interests

The author declares that she has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Not applicable. No participants were involved in this study.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Allan, K. (2001). *Natural language semantics*. Massachusetts: Blackwell Publishers.
- Bhatt, R. (2006). *Covert modality in non-finite contexts*. Berlin, Mouton de Gruyter. <https://doi.org/10.1515/9783110197341>
- Condoravdi, C. (2001). Temporal interpretation of modals: Modals for the present and for the past In D. Beaver, L. Casilas Martinez, B. Clark, & S. Kaufmany (Eds.), *The construction of meaning* (pp. 59-88). Stanford, CA: CSLI Publications.
- Gelderen, E. van (2010). *An introduction to the grammar of English*. Philadelphia, PA: John Benjamins Publishing Company. <https://doi.org/10.1075/z.153>
- Gillon, B. S. (2019). *Natural language semantics: Formation and evaluation*. Cambridge, MA: MIT Press.
- Grice, P. (1969). Utterer's meaning and intention. *The Philosophical Review*, 78(2), 147-177. Duke University Press on behalf of *Philosophical Review*. <https://doi.org/10.2307/2184179>
- Hacquard, V. (2006). *Aspects of modality*. (Unpublished doctoral thesis). Massachusetts, Massachusetts Institute of Technology.
- Hacquard, V. (2016). Modals: Meaning categories? In Blaszcak, J., Giannakidou, A., Klimek-Jankowska, D., & Migdalski, K (Eds.), *Mood, aspect, modality revisited: New answers to old questions* (pp. 45-74). Chicago, IL: The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226363660.003.0002>
- Hurford, J. R., Heasley, B., & Smith, M. B. (2007). *Semantics: A coursebook*. UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511841668>
- Kearns, K. (2011). *Semantics* (2nd ed.). New York, NY: Palgrave Macmillan. <https://doi.org/10.1007/978-0-230-35609-2>
- Kratzer, A. (1977). What "must" and "can" mean. *Linguistics and Philosophy*, 1(3), 337-355. <https://doi.org/10.1007/BF00353453>
- Löbner, S. (2014). *Understanding semantics*. USA: Routledge. <https://doi.org/10.4324/9780203777596>
- Lyons, J. (1977). *Semantics* (Vols. 1 & 2). UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620614>
- Ortega-Martín, M., García-Sierra, Ó., Ardoiz, A., Álvarez, J., Armenteros, J. C., & Alonso, A. (2023). *Linguistic ambiguity analysis in ChatGPT 4o*. arXiv. <https://doi.org/10.48550/arXiv.2302.06426>
- Poesio, M. (1996). Semantic ambiguity and perceived ambiguity. In K. van Deemter, & S. Peters (Eds.), *Semantic ambiguity and underspecification* (pp. 159-201). Stanford, CA: Center of the Study of Language and Information.
- Sennet, A. (2023). Ambiguity. *Stanford encyclopedia of philosophy*. Retrieved from <https://plato.stanford.edu/entries/ambiguity/>
- von Stechow, P., & Iatridou, S. (2008). How to say *ought* in foreign: The composition of weak necessity modals. In J. Guéron, & J. Lecarme (Eds.), *Time and modality. Studies in natural language and linguistic theory* (Vol. 75, pp. 115-141). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-8354-9_6

Appendix A: Linguistically Ambiguous Sentences

1. The old men and women left the room.
2. I watched the man with the telescope.
3. The chicken is ready to eat.
4. John and Jim are brothers I can't stand.
5. I shot an elephant in my pyjamas.
6. The man pushing the cart was tall.
7. The man eating the fish is sick.
8. They are hunting dogs.
9. I read the book on the balcony.
10. The professor said on Monday he would give an exam.
11. Going to the right is the right choice.
12. I saw a cat with one eye.
13. My mother and sister were sad after she shouted at her.
14. She knew a little Greek.
15. Superfluous hair remover.
16. John floated the boat between the rocks.
17. John ought to be at home by now.
18. The coin might come up heads.
19. You must eat a piece of cake.
20. Every woman squeezed a man.
21. Maria resolved the problem too.
22. The driver of my sister who lived there died.
23. The mother of my friend who bought the house left.
24. Everyone loves his mother.
25. John looked up when he came in.
26. Every cheerleader had shared a milkshake with Jones.
27. Every student knows two languages.
28. The company sent one representative to every meeting.
29. My mother and sister were sad after she shouted at her.

Appendix B: Ambiguous Sentences That Google Gemini Failed to Detect

1. They are hunting dogs.
2. I read the book on the balcony.
3. I saw a cat with one eye.
4. She knew a little Greek.
5. The coin might come up heads.
6. John ought to be at home by now.
7. The coin might come up heads.
8. Maria resolved the problem too.
9. The box of toys, which I bought, is expensive.
10. John floated the boat between the rocks.