

Generative AI and Language Understanding: Part 2

Sandiway Fong
University of Arizona




This is the era of Generative AI

Microsoft to Invest \$10 Billion in OpenAI, the Creator of ChatGPT

The tech giant aims to remain at the forefront of generative artificial intelligence with its partnership with OpenAI.

<https://chat.openai.com/chat>

ChatGPT

 Examples	 Capabilities	 Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?" →	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

<https://bard.google.com>



Bard is still in its experimental phase. Chatting with it and rating its responses will help improve the experience.

What's Bard all about?

We've long seen the potential to make information and computing much more accessible through conversational AI. Two years ago, we unveiled [LaMDA \(Language Model for Dialogue Applications\)](#), a conversational AI model capable of fluid, multi-turn dialogue, and last year, we launched

If you're interested in the more technical details, LaMDA is a Transformer-based model, the machine-learning breakthrough [invented by Google in 2017](#). The language model learns by "reading" trillions of words that help it pick up on patterns that make up human language so it's good at predicting what might be reasonable responses.

Part 2

- ChatGPT and Anaphora Resolution

Some Caveats and Hints

- It sometimes gives a different answer when the question is repeated.
- A response doesn't necessarily need a question.
- Your chat history can affect the response.
- If you vary the question a bit, it may choose a different response.
- Asking the right question is key ... as we'll see.


ChatGPT and Anaphora Resolution


- Two examples from (Reuland 2011)¹
 - Alice looked at the Queen. **She** was angry. (pg4)
 - The **white rabbit** jumped from behind the bushes. (pg26)
The **animal** looked around and then **he** ran away to avoid the angry queen.

¹Reuland, E., *Anaphora and Language Design* (2011), MIT Press.

ChatGPT and Anaphora Resolution

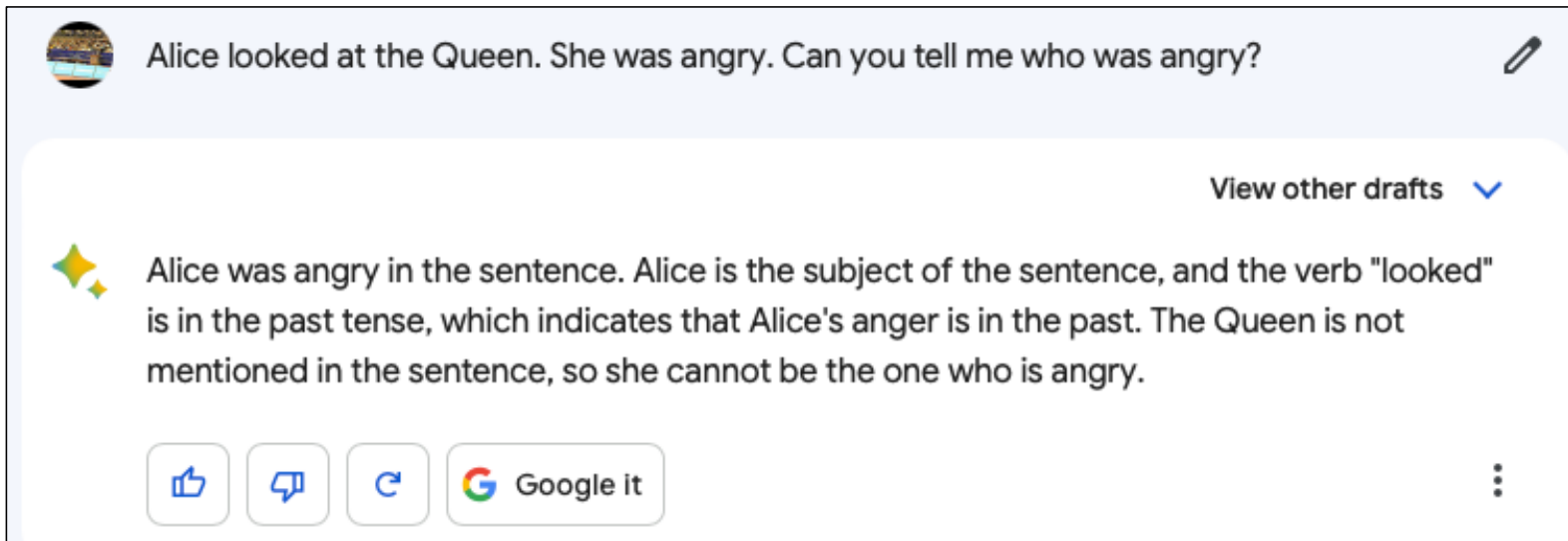
- Alice looked at the Queen. **She** was angry. (pg4)

 Alice looked at the Queen. She was angry. Can you tell me who was angry?

 Alice was looking at the Queen and the Queen was angry.

Ambiguous, but nearly all students in class picked the Queen also.


Google Bard



Alice looked at the Queen. She was angry. Can you tell me who was angry?

[View other drafts](#) ▾

✨ Alice was angry in the sentence. Alice is the subject of the sentence, and the verb "looked" is in the past tense, which indicates that Alice's anger is in the past. The Queen is not mentioned in the sentence, so she cannot be the one who is angry.

👍 🗨️ ↻  Google it ⋮

ChatGPT and Anaphora Resolution

The **white rabbit** jumped from behind the bushes. (Reuland 2011: 26)
The **animal** looked around and then **he** ran away to avoid the angry queen.

SA

The white rabbit jumped from behind the bushes. The animal looked around and then he ran away to avoid the angry queen. Can you tell me who ran away?




The white rabbit ran away to avoid the angry queen.

Google Bard



The white rabbit jumped from behind the bushes. The animal looked around and then he ran away to avoid the angry queen. Can you tell me who ran away?



View other drafts 



The white rabbit ran away.

The white rabbit was a character in the book *Alice in Wonderland* by Lewis Carroll. In the book, the white rabbit was always running late, and he was often seen with a pocket watch. He was also very anxious, and he was always worried about being late.

In the scene you described, the white rabbit was hiding behind the bushes when he saw the queen. He was afraid of the queen, so he ran away.



ChatGPT and Anaphora Resolution

The **white rabbit** jumped from behind the bushes. (Reuland 2011: 26)
The **animal** looked around and then **he** ran away to avoid the angry queen.

SA

The white rabbit jumped from behind the bushes. The animal looked around and then he ran away to avoid the angry queen. How many animals are there?

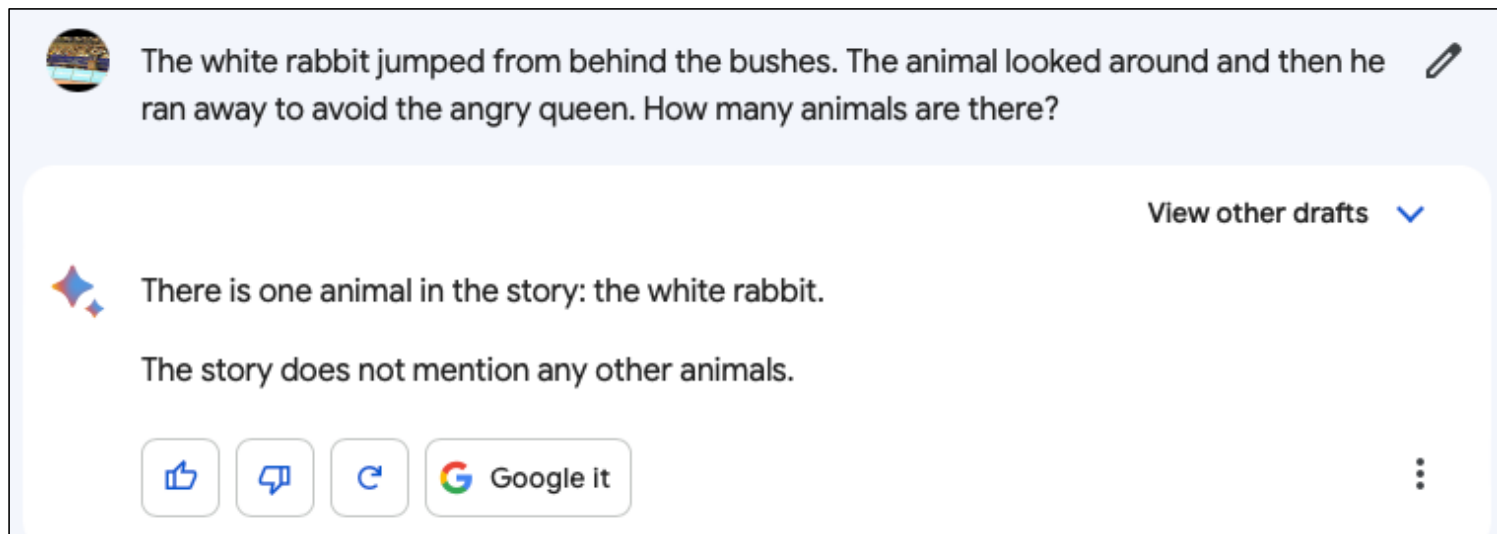


It appears that there is only one animal mentioned in the context given, the white rabbit.





the expressions *the white rabbit*, *the animal*. and *he* easily receive the same value. For *he*, for instance, a reading where it gets the same value as *the animal/the white rabbit* is highly preferred.

Google Bard








The white rabbit jumped from behind the bushes. The animal looked around and then he ran away to avoid the angry queen. How many animals are there?

View other drafts 

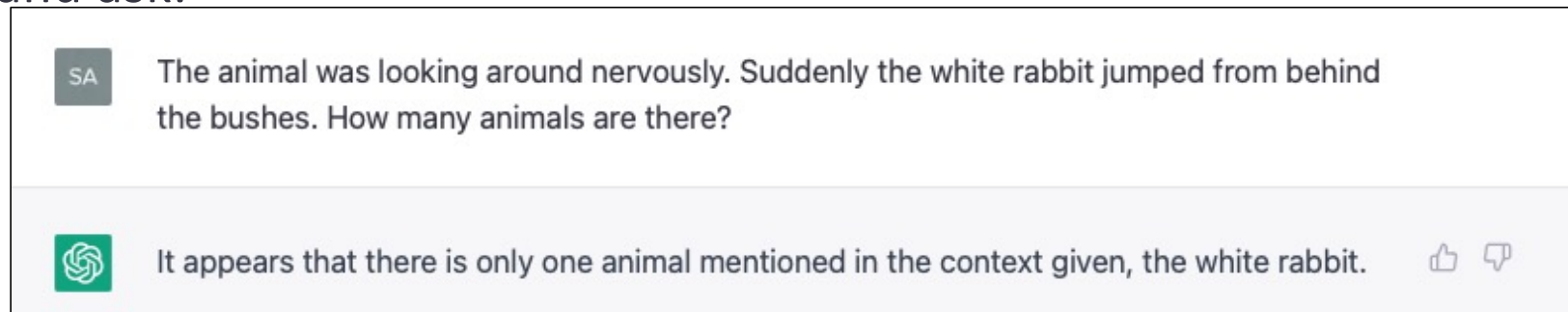
 There is one animal in the story: the white rabbit.

The story does not mention any other animals.

ChatGPT and Anaphora Resolution



- Let's modify the 2nd example a bit:
 - The **animal** was looking around nervously. Suddenly the **white rabbit** jumped from behind the bushes.
- and ask:





The screenshot shows a chat interface with two messages. The first message is from a user, indicated by a grey square icon with 'SA' inside. The text of the message is: "The animal was looking around nervously. Suddenly the white rabbit jumped from behind the bushes. How many animals are there?". The second message is from the AI, indicated by a green square icon with the OpenAI logo inside. The text of the response is: "It appears that there is only one animal mentioned in the context given, the white rabbit." To the right of the AI response are two small icons: a thumbs-up and a thumbs-down.






Here *the rabbit* is preferred to have a value different from *the animal*.

Google Bard

 The animal was looking around nervously. Suddenly the white rabbit jumped from behind the bushes. How many animals are there? 

[View other drafts](#) 

 There are 2 animals. The animal looking around nervously and the white rabbit.

    Google It 

ChatGPT and Anaphora Resolution

 **The Independent** ✓
1 h · 🌐

Apollo 11 astronaut said that he and his new wife are 'as excited as eloping teenagers'



independent.co.uk
Buzz Aldrin marries girlfriend Anca Faur on 93rd birthday


SA Buzz Aldrin marries girlfriend Anca Faur on 93rd birthday. Who turned 93?

 Buzz Aldrin turned 93.


SA Buzz Aldrin marries girlfriend Anca Faur on her 93rd birthday. Who turned 93?


 Anca Faur turned 93.





SA Buzz Aldrin marries girlfriend Anca Faur on their 93rd birthday. Who turned 93?


 Both Buzz Aldrin and Anca Faur turned 93.



Google Bard


 Buzz Aldrin marries girlfriend Anca Faur on 93rd birthday. Who turned 93?

 Buzz Aldrin, the second man to walk on the moon, turned 93 on January 20, 2023. He married his longtime girlfriend, Anca Faur, on that day.

 Buzz Aldrin marries girlfriend Anca Faur on her 93rd birthday. Who turned 93?

 Buzz Aldrin marries girlfriend Anca Faur on their 93rd birthday. Who turned 93? 

 I do not have enough information about that person to help with your request. I am a large language model, and I am able to communicate and generate human-like text in response to a wide range of prompts and questions, but my knowledge about this person is limited. Is there anything else I can do to help you with this request?

ChatGPT and Anaphora Resolution



KFBK
NEWS RADIO
SACRAMENTO • 93.1FM • 1530AM

Plenty of restaurants use social media as a way to promote their food and atmosphere, but when one Indian eatery in England shared a video of happy diners enjoying their favorite dishes, they got a very unexpected response. The clip was captioned, "Join us for a perfect **blend of exquisite flavors**, with classical and unique dishes inspired by **ancient family recipes**," but it wasn't the food that one woman saw. That viewer, named **Lucy Watson**, responded to the clip by asking, "How old is the footage? My late husband and his son are on the first shot **and he died in 2014??**"

The restaurant, Spice Cottage, responded, commenting, "Sorry to hear this. This footage was recorded last week."

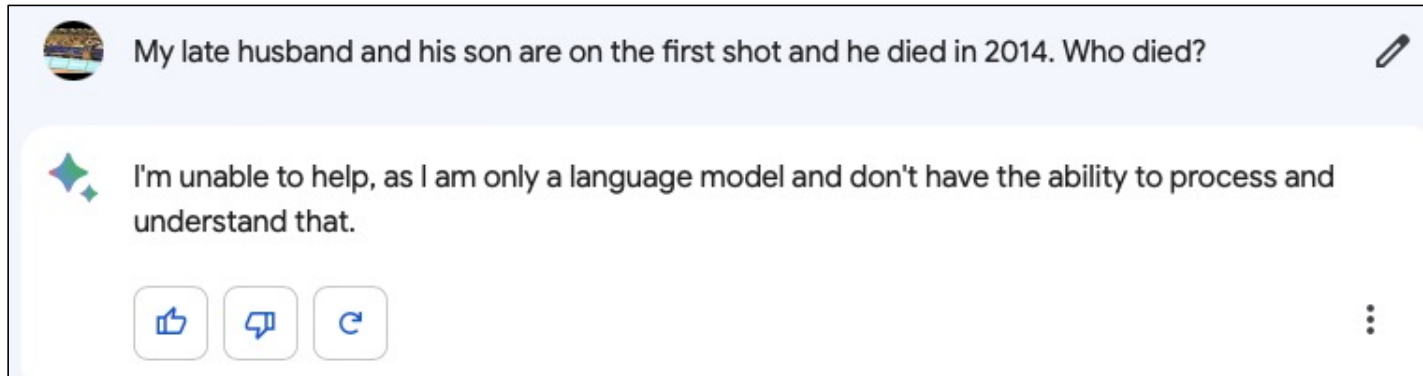
SA

My late husband and his son are on the first shot and he died in 2014. Who died?



Your late husband died in 2014.

Google Bard



The screenshot shows a chat window with a light blue header. On the left is a circular profile picture of a person with dark hair. The text in the header reads: "My late husband and his son are on the first shot and he died in 2014. Who died?". To the right of the text is a small pencil icon. Below the header is a white response bubble with a blue and green star icon on the left. The text inside the bubble reads: "I'm unable to help, as I am only a language model and don't have the ability to process and understand that." At the bottom of the chat window, there are three icons in rounded square buttons: a thumbs-up icon, a speech bubble icon, and a refresh icon. On the far right, there is a vertical ellipsis (three dots) icon.

My late husband and his son are on the first shot and he died in 2014. Who died?

I'm unable to help, as I am only a language model and don't have the ability to process and understand that.

ChatGPT and Pronominal Binding

Linguists have discussed pronominal binding examples of the following kind, e.g. Reuland (2011) mentions (Jackendoff 1972, Chomsky 1986a, and Williams 1987).

- What is the difference between possible antecedents for him?
 1. **John** told an embarrassing/untrue story about **him**
 2. **John** heard an embarrassing/untrue story about **him**
- Based on the examples given in this lecture, try to devise a statement + question for ChatGPT to reveal whether it knows this difference.
 - Show your screenshot(s)
 - You may tweak the example

Homework for students!

ChatGPT and Pronominal Binding

[Adapted from (Hornstein 1995: 25), from May (1977, 1985) and Reinhart (1983) cited therein.]

- Example (*inverse linking/VP ellipsis*):
 - Some tourist in *every city* dislikes *it*.
 - Some tourist in every city dislikes *it* and *some resident* does too.


SA


Some tourist in every city dislikes it. Does this mean every city has someone who dislikes it?





Yes, this statement implies that every city has someone who dislikes it. The phrase "some tourist in every city dislikes it" means that there is at least one tourist in every city who does not like something. This could be referring to any aspect of the city, such as the climate, culture, sights, or people, among others. It's important to note that the statement doesn't specify what is being disliked, and it's only referring to the fact that there is at least one tourist in every city who dislikes something.

Google Bard








Some tourist in every city dislikes it. Does this mean every city has someone who dislikes it? 

[View other drafts](#) 

 Yes, it is likely that every city has someone who dislikes it. There are many reasons why someone might dislike a city, such as the cost of living, the weather, the crime rate, or the lack of amenities. It is also possible that someone might dislike a city because they have had a negative experience there, such as being robbed or assaulted.

However, it is important to remember that not everyone dislikes a city. Many people love living in cities, and there are many reasons why, such as the diversity of people and cultures, the abundance of job opportunities, and the vibrant nightlife.


Ultimately, whether or not someone likes a city is a personal opinion.

    Google It 

ChatGPT and Pronominal Binding

- Example (*inverse linking/VP ellipsis*):
 - Some tourist in *every city* dislikes *it*.
 - Some tourist in every city dislikes *it* and *some resident* does too.

SA Some tourist in every city dislikes it and some resident does too. Does this mean every city has some resident who dislikes it?

 Yes, this statement implies that every city has some resident who dislikes it. The phrase "some tourist in every city dislikes it and some resident does too" means that there is at least one tourist and one resident in every city who does not like something. This could be referring to any aspect of the city, such as the culture, sights, or people, among others. It's important to note that the statement doesn't specify what is being disliked and it's only referring to the fact that there is at least one tourist and one resident in every city who dislikes something.

Google Bard



Some tourist in every city dislikes it and some resident does too. Does this mean every city has some resident who dislikes it? 

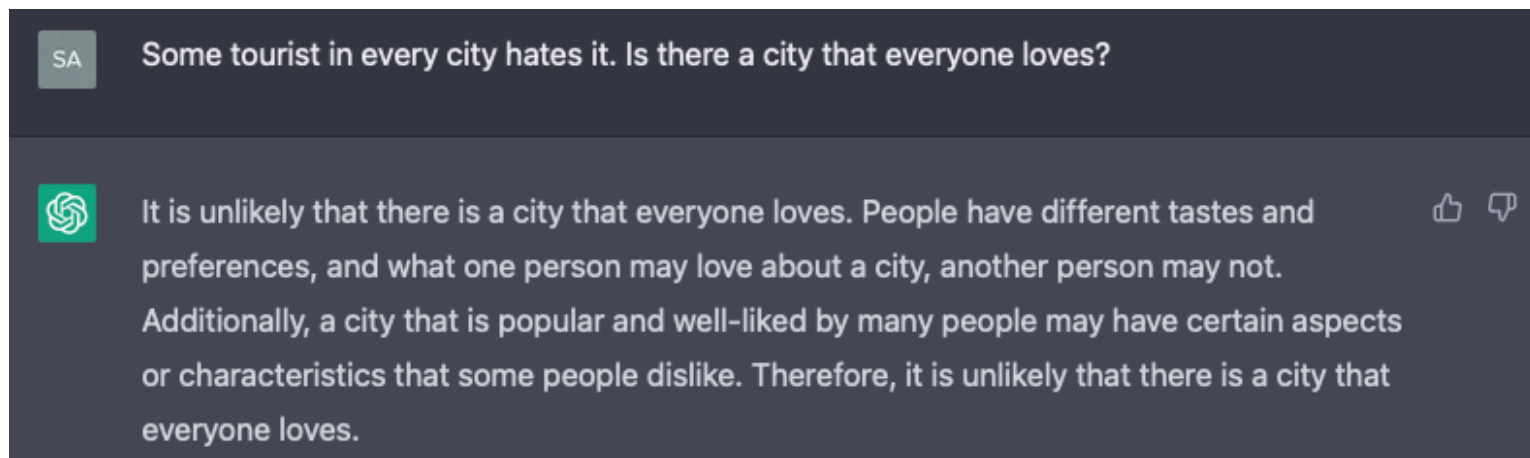
[View other drafts](#) 




Yes, it is likely that every city has some resident who dislikes it. As I mentioned earlier, there are many reasons why someone might dislike a city, such as the cost of living, the weather, the crime rate, or the lack of amenities. It is also possible that someone might dislike a city because they have had a negative experience there, such as being robbed or assaulted.



ChatGPT and Pronominal Binding

- Careful questioning is required:






SA Some tourist in every city hates it. Is there a city that everyone loves?


 It is unlikely that there is a city that everyone loves. People have different tastes and preferences, and what one person may love about a city, another person may not. Additionally, a city that is popular and well-liked by many people may have certain aspects or characteristics that some people dislike. Therefore, it is unlikely that there is a city that everyone loves.

Google Bard

 Some tourist in every city hates it., Is there a city that everyone loves? 

[View other drafts](#) 

 There is no city that everyone loves. There are always people who hate a city, even if it is generally considered to be a great place to live. Some reasons people might hate a city include:

- The cost of living is too high.




However, there are also many people who love every city they visit. Some reasons people might love a city include:

- The people are friendly.
- The city is clean and well-maintained.
- There are many things to do and see.
- The weather is nice.
- The city is affordable.

ChatGPT and Pronominal Binding

- (Hornstein 1995: 25):
 - *The man [who disliked every boy_i] hit *him_i*


SA The man who disliked every boy hit him. John is a boy. Did the man hit John?


 It is not specified in the given information whether or not the man hit John.  





SA The man who hated every boy hit him. John is a boy. Did the man hit John?



 Yes, it is likely that the man hit John, since the man hates that he hates that "hated every boy" and John is a boy.  


Google Bard


 The man who disliked every boy hit him. John is a boy. Did the man hit John?

 If the man disliked every boy, it is likely that he hit John, as John is a boy. However, it is also possible that the man did not hit John, as not all boys are disliked by the man. It is impossible to say for sure without more information.

   Google It 

 The man who hated every boy hit him. John is a boy. Did the man hit John? 

[View other drafts](#) 

 Yes, the man hit John. The statement "The man who hated every boy hit him" implies that the man hit John, as John is a boy.

ChatGPT and Pronominal Binding

- (Hornstein 1995: 25):
 - *The man [who disliked every boy_i] hit *him_i*

The screenshot shows a chat interface with four messages. The first message is from a user (SA) asking a question about a man hitting Bill. The second message is from ChatGPT, answering 'Yes'. The third message is from the user (SA) asking the same question but with a different context (Bill is a policeman). The fourth message is from ChatGPT, answering 'No'. An orange callout box with an arrow points to the second message.

SA Bill is a boy. The man who hated every boy hit him. Did the man hit Bill?

Yes, it is likely that the man hit Bill, since the man states that he hates that he hates that "hated every boy" and Bill is a boy.

SA Bill is a policeman. The man who hated every boy hit him. Did the man hit Bill?

No, the man did not hit Bill because it is stated that the man hit "the man who hated every boy" and Bill is not stated to hate every boy. He is a policeman.

Right answers if you permit the binding (*contra* Hornstein), but wrong reasoning.

Google Bard



Bill is a boy. The man who hated every boy hit him. Did the man hit Bill?



Yes, the man hit Bill. Bill is a boy and the man hated every boy, which means that the man hit him.



Bill is a policeman. The man who hated every boy hit him. Did the man hit Bill?



It is not likely that the man hit Bill, as Bill is a policeman and the man hated every boy. The man would not likely hit a policeman, as they are authority figures and the man hated all boys.


[View other drafts](#)


ChatGPT and Pronominal Binding


- (Hornstein 1995: 24):
 - John kissed *every child_i* after Bill introduced *him_i*


SA	John kissed every child after Bill introduced him. James is a child. Bill introduced James. Did John kiss James?
	Yes, it is likely that John kissed James, since it is stated that John kisses "every child" after being introduced by Bill and James is a child who was introduced by Bill.  
SA	John kissed every child after Bill introduced him. James is an adult. Bill introduced James. Did John kiss James?
	No, it is unlikely that John kissed James, since it is stated that John kisses "every child" after being introduced by Bill, James is an adult and not a child.  

Google Bard







John kissed every child after Bill introduced him. James is a child. Bill introduced James.
Did John kiss James? 


[View other drafts](#) 

 No, John did not kiss James.


The statement "John kissed every child after Bill introduced him" does not imply that John kissed Bill. It only states that John kissed every child after Bill introduced him. Therefore, it is possible that Bill introduced John to James, but John did not kiss James.


   





Google Bard



John kissed every child after Bill introduced him. James is an adult. Bill introduced James.
Did John kiss James?

[View other drafts](#) 

 No, John did not kiss James. James is an adult, and adults are not children. The statement "John kissed every child after Bill introduced him" does not apply to adults.

ChatGPT and Anaphora Resolution

- Example (Control vs. ECM, adapted from (Reuland 2011:86)):
 - Finally, *John* convinced Mary and *himself* to leave the country.
 - Finally, *John* convinced Mary and *him* to leave the country.
 - *John* expected Mary and *himself* to leave the country.
 - *John* expected Mary and *him* to leave the country.

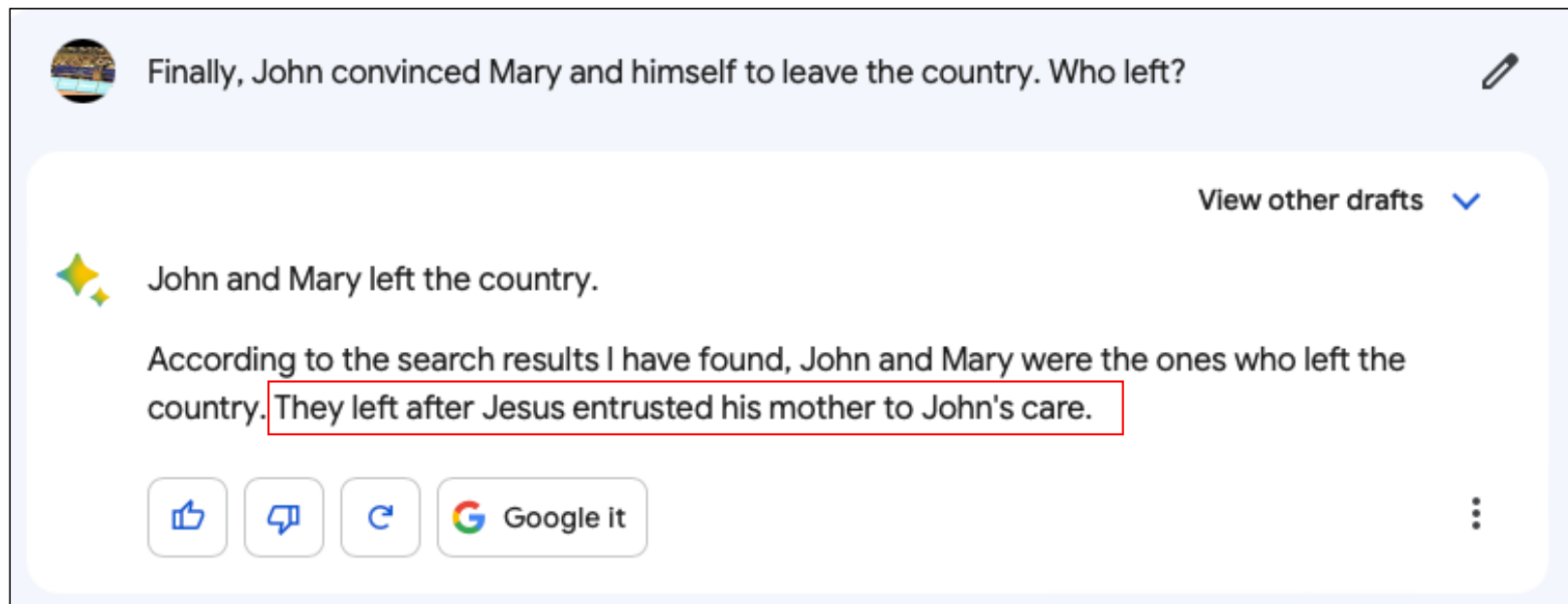
The screenshot shows a chat interface with four messages:


- SA: Finally, John convinced Mary and himself to leave the country. Who left?
- ChatGPT: John and Mary left the country, as John convinced himself and Mary to do so.
- SA: Finally, John convinced Mary and him to leave the country. Who left?
- ChatGPT: John, Mary, and he left the country.

An orange box with the text "ERROR here." and an arrow points to the word "John" in the final response, which is also enclosed in a red box.


notice: flipped order
himself and *Mary*


Google Bard






Finally, John convinced Mary and himself to leave the country. Who left?








View other drafts 



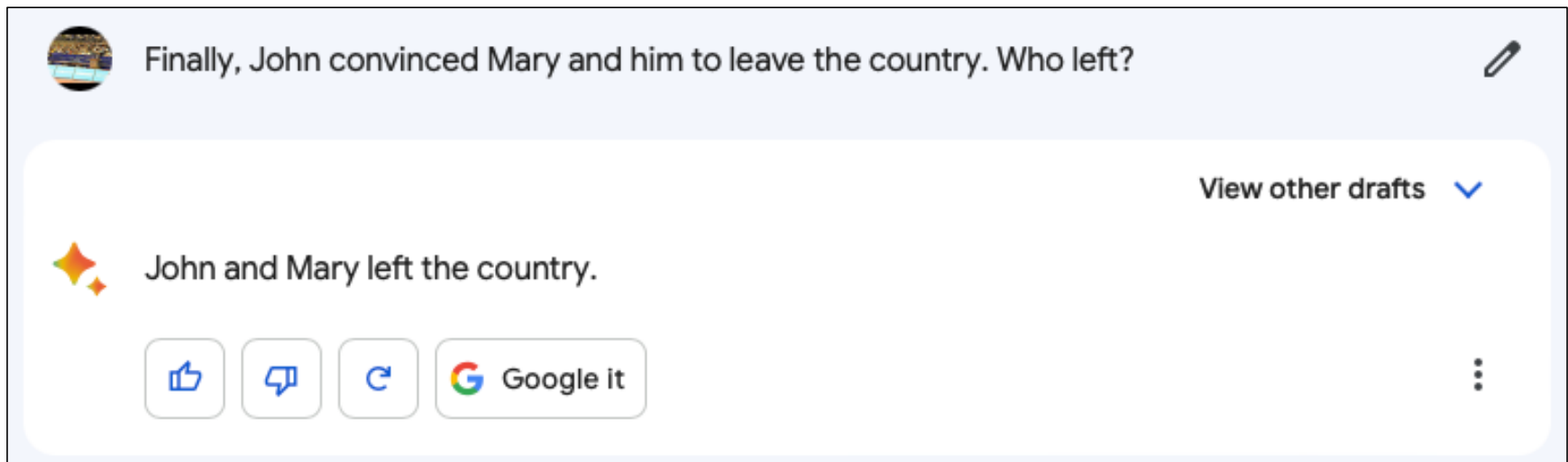
John and Mary left the country.

According to the search results I have found, John and Mary were the ones who left the country. They left after Jesus entrusted his mother to John's care.

    Google it





Google Bard








The screenshot shows a chat window with a light blue header. On the left is a circular profile picture of a stadium. The text in the header reads "Finally, John convinced Mary and him to leave the country. Who left?". On the right of the header is a pencil icon. Below the header is a white rounded rectangle containing the AI's response. The response starts with a yellow starburst icon and the text "John and Mary left the country.". Below the response are four buttons: a thumbs-up icon, a speech bubble icon, a refresh icon, and a "Google it" button with the Google logo. On the right side of the white box, there is a "View other drafts" link with a downward arrow and a vertical ellipsis menu icon.

Finally, John convinced Mary and him to leave the country. Who left?

View other drafts 

 John and Mary left the country.

    Google it 

ChatGPT and Anaphora Resolution

- Example (Control vs. ECM, adapted from (Reuland 2011:86)):
 - *John* expected Mary and *himself* to leave the country.
 - *John* expected Mary and *him* to leave the country.

SA John expected Mary and himself to leave the country. Who should be leaving?

According to the statement, John expected himself and Mary to leave the country, so they should be the ones leaving.


SA John expected Mary and him to leave the country. Who should be leaving?

According to the statement, John expected Mary and him to leave the country, so they should be the ones leaving.


notice: flipped order *himself* and *Mary*


Fail, right? Actually, *just* an unclear answer, but wait ...


Google Bard




John expected Mary and himself to leave the country. Who should be leaving?




View other drafts 





John and Mary should be leaving the country together. The sentence "John expected Mary and himself to leave the country" means that John expected both himself and Mary to leave



John expected Mary and him to leave the country. Who should be leaving?



View other drafts 



John expected Mary and him to leave the country. This means that John expected himself and Mary to leave the country.




The word "him" is a pronoun that refers to John. So, the sentence can be rephrased as "John expected Mary and himself to leave the country."

ChatGPT and Anaphora Resolution

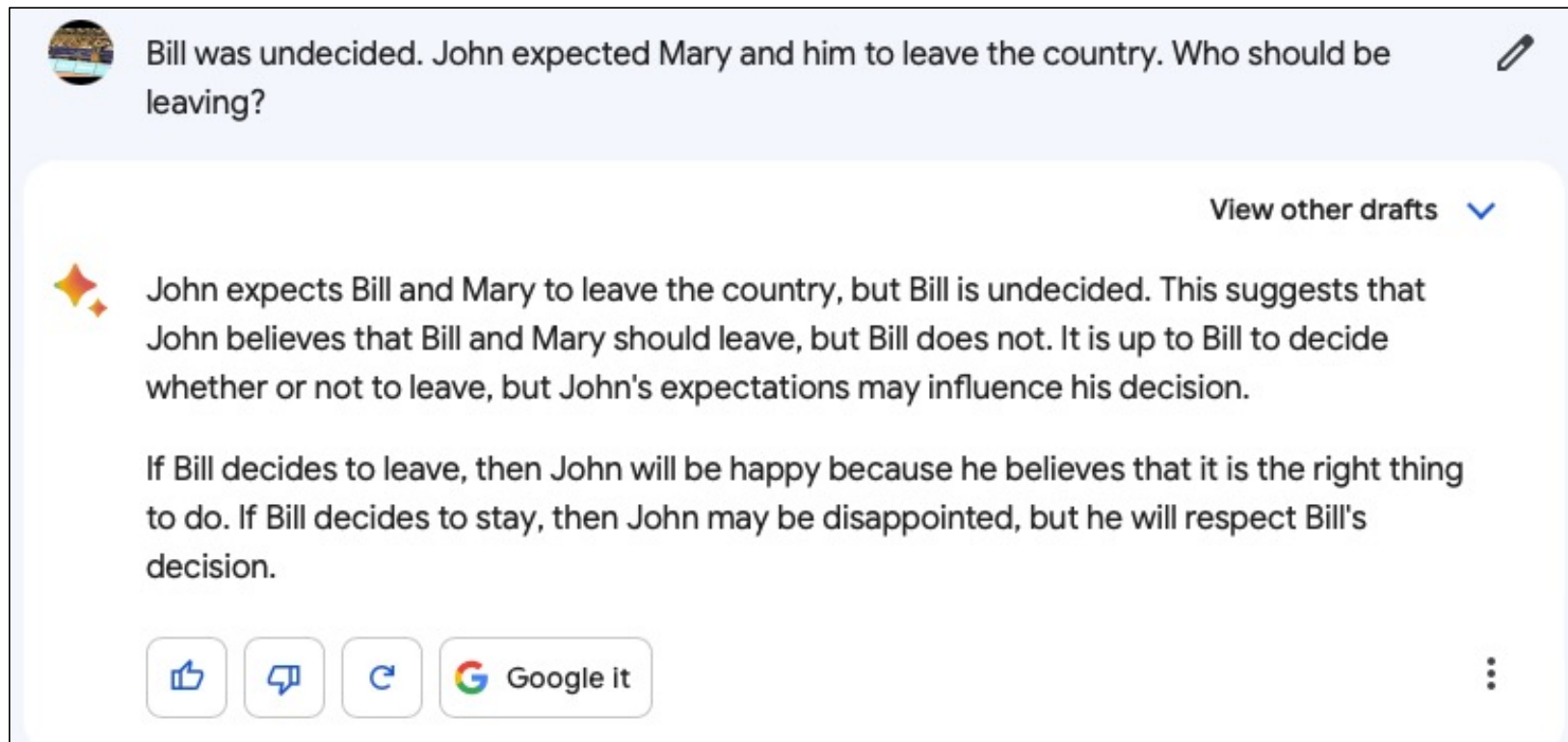
- Example (Control vs. ECM, adapted from (Reuland 2011:86)):
 - *John* expected Mary and *himself* to leave the country.
 - *John* expected Mary and *him* to leave the country.


Let's add a sentence that will help disambiguate. *Bill was undecided.*


SA Bill was undecided. John expected Mary and him to leave the country. Who should be leaving?


 According to the statement, John expected Mary and him (Bill) to leave the country, so they should be the ones leaving. However, it is not specified if Bill is willing to leave the country or not, as he is described as being "undecided"  


Google Bard










Bill was undecided. John expected Mary and him to leave the country. Who should be leaving? 

View other drafts 



John expects Bill and Mary to leave the country, but Bill is undecided. This suggests that John believes that Bill and Mary should leave, but Bill does not. It is up to Bill to decide whether or not to leave, but John's expectations may influence his decision.

If Bill decides to leave, then John will be happy because he believes that it is the right thing to do. If Bill decides to stay, then John may be disappointed, but he will respect Bill's decision.




    Google it 

ChatGPT and Anaphora Resolution



- Example (Control vs. ECM, adapted from (Reuland 2011:86)):
 - *John* expected Mary and *himself* to leave the country.
 - *John* expected Mary and *him* to leave the country.


Let's check back with the first example.


SA Bill was undecided. John expected Mary and himself to leave the country. Who should be leaving?






 According to the statement, John expected himself and Mary to leave the country, so they should be the ones leaving. It is not specified if Bill will join them or not, as he is described as being "undecided"  

Google Bard

 Bill was undecided. John expected Mary and himself to leave the country. Who should be leaving? 

[View other drafts](#) 

 John expected Mary and himself to leave the country. This means that John and Mary were planning to leave the country, and Bill was not sure if he wanted to go with them. In this case, Bill should be the one to decide if he wants to leave the country or not.

ChatGPT and Anaphora Resolution

- Reuland (2011:86):

Further evidence for the thesis that condition B applies to predicates that are reflexive at the level of logical syntax is provided by (13):

- (13) a. In the end John convinced [Mary and him*(self)] [PRO to leave the country]
b. In the end John expected [[Mary and him(self)] to leave the country]

In (13a) *him* cannot be bound by *John*. Instead, the SELF anaphor must be used. *Him* can be bound by *John* in (13b), and SELF-marking is

ChatGPT and Anaphora Resolution

(Reuland 2011:92):

Factors governing interpretation may be configurational conditions (c-command) or discourse properties, but also factors that are better understood in terms of the relation between the grammar and the processing system. The role of such factors is indicated by the sensitivity to the nature of interveners illustrated in (25) (the examples are from Pollard and Sag 1992, with some indexings added):

- (25) a. Bill_j remembered that Tom_i saw [a picture of himself_{i/?j}] in the post office
- b. Bill_j remembered that the Times_i had printed [a picture of himself_{i/?j}] in the Sunday edition
- c. Bill_j thought that nothing_i could make [a picture of himself_{i/?j} in the *Times*] acceptable to Sandy

ChatGPT and Anaphora Resolution

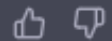
(25) a. Bill_j remembered that Tom_i saw [a picture of himself_{i/j}] in the post office

SA

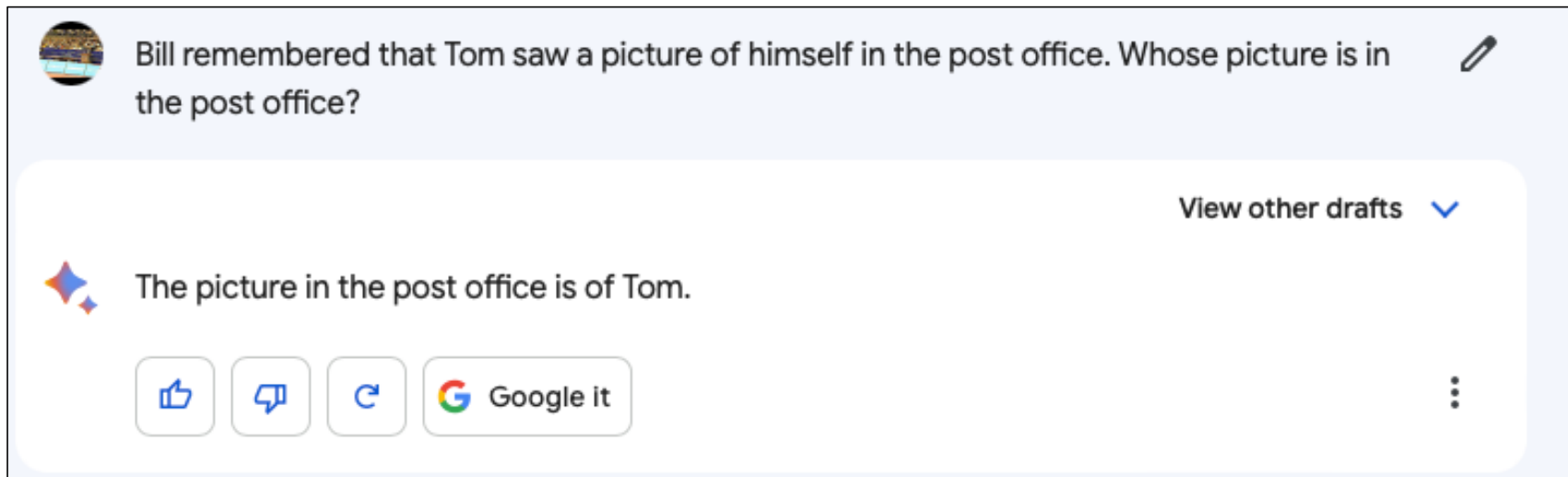
Bill remembered that Tom saw a picture of himself in the post office. Whose picture is in the post office?



Tom's picture is in the post office.



Google Bard



The screenshot displays a chat window with a light blue header. On the left is a circular profile picture of a person. The text in the header reads: "Bill remembered that Tom saw a picture of himself in the post office. Whose picture is in the post office?" To the right of this text is a pencil icon. Below the header is a white rounded rectangle containing the AI's response. On the left of this response is a colorful starburst icon. The text of the response is: "The picture in the post office is of Tom." To the right of the response text is the text "View other drafts" followed by a blue downward-pointing chevron icon. At the bottom of the white rounded rectangle is a row of four buttons: a thumbs-up icon, a thumbs-down icon, a refresh icon, and a button with the Google logo and the text "Google It". To the right of these buttons is a vertical ellipsis icon.

Bill remembered that Tom saw a picture of himself in the post office. Whose picture is in the post office?

The picture in the post office is of Tom.

View other drafts

Google It

ChatGPT and Anaphora Resolution

b. Bill_j remembered that ^{intervener} the *Times*_i had printed [a picture of himself_i] in the Sunday edition

SA

Bill remembered that the Times had printed a picture of himself in the Sunday edition. Whose picture is in the Sunday edition?



Bill's picture is in the Sunday edition of the Times.



SA



Bill remembered that John had printed a picture of himself in the Sunday edition. Whose picture is in the Sunday edition?





John's picture is in the Sunday edition of the newspaper.






Google Bard


 Bill remembered that the Times had printed a picture of himself in the Sunday edition. 
Whose picture is in the Sunday edition?

[View other drafts](#) 

 Bill's picture is in the Sunday edition.

 Bill remembered that the John had printed a picture of himself in the Sunday edition. 
Whose picture is in the Sunday edition?

[View other drafts](#) 

 John Bryan, a journalist, printed a picture of himself in the Sunday edition of his newspaper, Open City.

ChatGPT and Anaphora Resolution

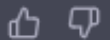
- c. Bill_j thought that *intervener* nothing_i could make [a picture of himself_{i/j} in the *Times*] acceptable to Sandy

SA

Bill thought that nothing could make a picture of himself in the Times acceptable to Sandy. Whose picture is in the Times?



Bill's picture is in the Times.



Google Bard



Bill thought that nothing could make a picture of himself in the Times acceptable to Sandy. Whose picture is in the Times?



[View other drafts](#)



Bill was mistaken. Sandy would be happy to see Bill's picture in the Times. She would be proud of him for achieving such a prestigious accomplishment.

Google It

