



Universidad
Zaragoza

TRABAJO FIN DE GRADO

Procesamiento de Ambigüedades en Modelos de Lenguaje

AUTORA: EMMA ANGLÉS-HERRERO

DIRECTORES: JOSÉ LUIS MENDÍVIL GIRÓ Y JORGE GRACIA DEL RÍO

FACULTAD DE FILOSOFÍA Y LETRAS
31 DE MAYO DE 2024

Resumen: La ambigüedad es un rasgo inherente de las lenguas naturales. Es por ello por lo que en este trabajo hemos puesto a prueba al modelo ChatGPT, para comprobar qué sensibilidad tiene a la hora de desentrañar ambigüedades de distintos tipos. Hemos elaborado un corpus, siguiendo la anotación de *natural language inference*, con tres tipos de ambigüedades: léxica, sintáctica y semántica. Haciendo distintas pruebas de *prompting* hemos podido comprobar que el modelo tiene capacidades para realizar inferencias, pero los resultados están alejados de los hablantes nativos de español.

Palabras clave: PLN, ambigüedad, Modelos de lenguaje, *natural language inference*, ChatGPT, desambiguación.

Abstract: Ambiguity is an inherent feature of natural languages, which is why in this work we have tested the ChatGPT model to see how sensitive it is to unravel ambiguities of different types. We have elaborated a corpus, following the natural language inference annotation, with three types of ambiguities: lexical, syntactic and semantic. By performing different prompting tests we have been able to verify that the model is capable of making inferences, but the results are far from what we could perform with Spanish speakers.

Key words: NLP, ambiguity, Language models, natural language inference, ChatGPT, disambiguation.

ÍNDICE

1.	INTRODUCCIÓN: CONTEXTO Y OBJETIVOS DEL TRABAJO	4
2.	LA CREACIÓN DEL CORPUS	9
3.	EXPERIMENTO	13
4.	RESULTADOS	16
4.1.	AMBIGÜEDAD LÉXICA	17
4.2.	AMBIGÜEDAD SINTÁCTICA.....	21
4.3.	AMBIGÜEDAD SEMÁNTICA	26
5.	CONCLUSIONES	31
6.	BIBLIOGRAFÍA	34
7.	ANEXOS.....	36

1. INTRODUCCIÓN: CONTEXTO Y OBJETIVOS DEL TRABAJO

El *procesamiento del lenguaje natural*¹ (PLN) es una disciplina cuyo objetivo es permitir que algunos programas informáticos interpreten de forma adecuada el lenguaje natural para llevar a cabo tareas de forma satisfactoria (Chowdhary, 2012). Algunas aplicaciones de esta disciplina incluyen clasificación de texto, traducción automática, extracción de información y otras muchas tareas. Sin embargo, esto plantea numerosos retos a diferentes niveles, ya que se necesita una gran cantidad de información lingüística del tipo léxico, sintáctico o semántico e información, que va más allá de las reglas y la formación de constituyentes, del tipo pragmático o inferencial. Abordar estos desafíos implica una comprensión profunda de la diversidad y complejidad del lenguaje humano, así como la capacidad de los sistemas para adaptarse a las variaciones y ambigüedades inherentes al mismo. Además, algunos sistemas de PLN se basan en probabilidad, lo que significa que el tamaño de la muestra de datos que se adhiera al modelo debe ser considerable para garantizar resultados confiables y precisos (Bates, Bobrow y Weischedel, 2006). Por lo tanto, el desarrollo y la mejora continua de los algoritmos y modelos de PLN son esenciales para avanzar en este campo. Esto hace que las aplicaciones de procesamiento de lenguaje natural sean más eficientes y efectivas en una amplia variedad de contextos y dominios de aplicación.

«La ambigüedad es la capacidad que tienen algunos enunciados o unidades léxicas de admitir más de una interpretación» (Icaraz Varó y Martínez Linares, 2004: 47). Este fenómeno fuerza a los receptores a tener que decidir en cada contexto qué significado e intención está activando el emisor (Piantadosi, Tily y Gibson, 2012). A pesar de ello, este hecho lingüístico no nos resulta extraño como hablantes de lenguas naturales, ya que como han señalado numerosos autores como Pinker y Bloom (1990: 772): «La ambigüedad es una consecuencia directa de la economía de formas superficiales: Si se reutiliza una palabra antigua para darle un nuevo significado, en lugar de inventar una nueva, se produce una ambigüedad cada vez que se utiliza la raíz»². Lo que demuestra la afirmación de la economía del lenguaje es la riqueza de usos y contextos en que se pueden utilizar diferentes palabras o estructuras sin que tengamos problemas en la interpretación de las oraciones o unidades léxicas. Este fenómeno, lejos de ser un obstáculo insalvable,

¹ También conocido en inglés como *Natural Language Processing*.

² «Ambiguity is a direct consequence of economy in surface forms: Reuse an old word for a new meaning, rather than inventing a new one, and you get an ambiguity whenever the stem is used.»

puede verse como un recurso comunicativo que ofrece flexibilidad y adaptabilidad en la interacción verbal. Esto permite a los hablantes expresar una amplia gama de matices y significados de manera concisa y eficaz.

El proceso por el que se resuelve este fenómeno se conoce como *desambiguación*, término de gran importancia para el presente trabajo. Podemos tener varios tipos de ambigüedades, como puede ser estructural, provocada por diferentes adjunciones entre los constituyentes (Icaraz Varó y Martínez Linares, 2004), o léxica provocada por diferentes fenómenos lingüísticos como la polisemia, la homonimia, etc. (Icaraz Varó y Martínez Linares, 2004). Como podemos ver, la ambigüedad es un fenómeno semántico intrínseco a la lengua, y sobre esta manifestación lingüística vamos a construir el presente trabajo.

En el ámbito del Procesamiento de Lenguaje Natural, es imposible no hacer referencia a los modelos de lenguaje (LMs)³. Estos son sistemas sofisticados que se encargan de asignar probabilidades a secuencias de palabras, permitiendo la generación de texto que parece humano y la clasificación del texto en categorías (Jurafsky y Martin, 2024).

Dentro de estos modelos de lenguaje, existen numerosas variedades y tipos. Desde algunos modelos más simples basados en los n-gramas hasta aquellos más complejos basados en arquitecturas de “transformers”. Los n-gramas se utilizan para predecir la siguiente palabra en una secuencia dependiendo de una ventana concreta de palabras anteriores. La probabilidad dependerá de la cantidad de veces que aparece un elemento junto a otro en un corpus de datos (Jurafsky y Martin, 2024). Los modelos basados en “transformers”, una arquitectura de redes neuronales que utilizan el llamado ‘mecanismo de autoatención’ (Vaswani, et al., 2017) que ha revolucionado el campo del PLN, también son conocidos como *autoregressive language models*. El mecanismo de autoatención puede ser concebido como una forma de construir representaciones contextuales del significado de una palabra que integra la información de las palabras que la rodean, ayudando al modelo a aprender cómo las palabras se relacionan con las otras en grandes extensiones de texto⁴ (Jurafsky y Martin, 2024: 214).

³ También conocidos como *Language Models* de ahí sus siglas LMs

⁴ Traducción directa de la referencia original en inglés

Aquellos modelos basados en esta arquitectura que se utilizan para generar texto, imagen, código, etc. se conocen como modelos generativos. La generación de este texto está condicionada por un *input* que será un trozo de texto al que llamaremos *prompt* (Jurafsky y Martin, 2024), pero por no hacer demasiado extensa esta introducción, dejaremos la explicación de la tarea de *prompting* para el epígrafe 3 de este trabajo. Algunos de los modelos de lenguaje más conocidos (en sus diferentes versiones) son: BERT, GPT o Llama. GPT es la base de aplicaciones muy utilizadas como ChatGPT⁵ o Copilot⁶.

ChatGPT es un *chatbot* o agente conversacional creado por OpenAI. Está basado en un modelo generativo por lo que puede generar texto de forma coherente dando respuesta a las entradas de texto que proporcionan los usuarios (OpenAI). Este modelo tiene en su entrenamiento grandes cantidades de datos textuales. A pesar de la gran utilidad que supone este modelo, debemos tener en cuenta que puede cometer errores o que esté sesgado a causa de los textos utilizados en su entrenamiento, como explican los creadores en su página web⁷. Es este modelo el que pondremos a prueba en nuestro trabajo. Hemos utilizado la versión GPT-3.5. Como se explica en la documentación pública de OpenAI, este modelo tiene una ventana de contexto de 4.096 *tokens* y es capaz de devolver un máximo de 4.096 *tokens* en su respuesta. Los datos de entrenamiento contienen información hasta septiembre de 2021. Este modelo ha sido entrenado para entender lenguaje natural y formal.

Los *tokens* son secuencias de caracteres producidas por una segmentación dentro del ‘análisis lingüístico’ que realizan los modelos de lenguaje, proceso se conoce como *tokenización* (Jurafsky y Martin, 2024). Este proceso es un paso más en la normalización del texto para tener una uniformidad y estandarización, y facilitar el posterior procesamiento. Los *tokens* no siempre se corresponden con palabras, ya que las segmentaciones pueden cambiar dependiendo del tipo de tarea que queremos realizar o del enunciado en el que nos encontremos (Mielke, 2021). Por ejemplo, si tenemos una frase como: ‘los gatos negros’, el *tokenizador* realizará una segmentación palabra por palabra: [‘los’, ‘gatos’, ‘negros’], en cambio si encontramos un enunciado con un

⁵ <https://chatgpt.com/>

⁶ <https://copilot.microsoft.com>

⁷ <https://platform.openai.com/docs>

significado unitario como ‘buenas noches’, el *tokenizador* lo tomará como un segmento único: [‘buenas noches’].

Este trabajo se concibe como un pequeño acercamiento al mundo del PLN. No con el propósito de usar las herramientas que nos ofrecen los modelos de lenguaje de forma baladí, sino el de intentar experimentar con técnicas propias del ámbito del PLN. Para este contacto, hemos decidido tomar como punto de partida un artículo titulado *We’re afraid language models aren’t modeling ambiguity* de Liu, Alisa, et al. (2023). El objetivo principal de nuestro trabajo es la familiarización con la creación de un corpus y las distintas pruebas que se pueden hacer para descubrir información sobre los modelos de lenguaje, mediante la réplica de uno de los experimentos propuesto en dicho artículo.

Los autores del artículo al que aludimos en el párrafo anterior tienen la intención de descubrir qué entendimiento de la ambigüedad pueden llegar a adquirir y detectar los modelos de lenguaje mediante diferentes pruebas. En concreto estas tres: generación de desambiguaciones, reconocimiento de desambiguaciones y modelaje de continuaciones específicas de la interpretación. Asimismo, ponen a prueba diferentes modelos de lenguaje para evaluar su rendimiento: LLaMa, GPT-3, FLAN-T5, InstructGPT, ChatGPT y GPT-4 mediante un *data set* formado por ambigüedades llamado AMBIENT creado por los autores del artículo (Liu, Alisa, et al., 2023). Además, para mostrar la importancia de tener unos modelos de lenguaje que sean sensibles en el reconocimiento de ambigüedades han utilizado un posible caso de uso real: la detección de enunciados engañosos, a causa de la vaguedad o ambigüedad, pronunciados por políticos para desinformar o desviar la atención del oyente.

Este estudio de la ambigüedad puede resultar de gran interés, ya que podemos poner a prueba las capacidades lingüísticas y extralingüísticas de ciertos modelos. Al mismo tiempo, como bien señalan Liu, Alisa, et al., (2023), muchas veces este fenómeno ha quedado relegado en el estudio del procesamiento del lenguaje natural por la complejidad que supone. Este problema muchas veces pasa desapercibido porque una persona no considera cada posibilidad cuando entiende una frase, pero los modelos o programas informáticos deben considerar cada una de ellas (Allen, 2024).

Lo que nosotros queremos conseguir con este trabajo es realizar una réplica a menor escala del primero de los experimentos: generación de desambiguaciones. Para ello se ha abordado la creación de un pequeño corpus de datos adaptado al español de la misma

manera que el artículo original, siguiendo la anotación de *Natural Language Inference* que explicaremos más adelante.

Es crucial no olvidar el análisis lingüístico que podemos aportar a los resultados. Aunque el artículo que hemos elegido como modelo para nuestro estudio se centra en la verificación de datos y números, nosotros podemos ir un paso más allá. Al incluir un análisis lingüístico en nuestros hallazgos, añadimos una capa adicional de interpretación, permitiendo una comprensión más profunda y matizada de los datos e intentando mostrar en qué casos el modelo de lenguaje puede ser más sensible a hacer una correcta desambiguación y qué casos lo es menos.

2. LA CREACIÓN DEL CORPUS

Para poder replicar el experimento del que hablaremos en el siguiente apartado, nos hemos encontrado con una gran limitación: la falta de un *corpus* de ambigüedades específico con el formato de *Natural Language Inference* (NLI) en español. El único artículo que hace una recopilación de oraciones en español con este formato no tiene un tratamiento específico para las ambigüedades (Kovarchev y Taulé, 2022). Por ello hemos optado por una elaboración manual. El NLI es una tarea automatizada que se ha utilizado para determinar la relación que hay entre dos textos. Puede haber tres tipos de relaciones: consecuencia lógica (*entailment*), si una hipótesis es verdadera dada una premisa; contradicción (*contradiction*), si una hipótesis es falsa dada una premisa y neutral (*neutral*), si no se puede determinar la veracidad de una hipótesis a partir de una premisa (Kovarchev y Taulé, 2022); (MacCartney, 2009). Una forma de ver cómo se presenta esta tarea es mediante el siguiente ejemplo⁸:

Premisa: ‘*John goes to work every day with a car.*’

Hipótesis: ‘*John has a job.*’

Relación: *entailment*. Podemos determinar que la relación es verdadera porque en la premisa ya se nos plantea que John tiene un trabajo.

Premisa: ‘*John goes to work every day with a car.*’

Hipótesis: ‘*John takes the bus to go to work.*’

Relación: *contradiction*. Podemos determinar que la relación es falsa porque en la premisa se nos presenta que va en coche y no en autobús como afirma la hipótesis.

Premisa: ‘*John goes to work every day with a car.*’

Hipótesis: ‘*John has a Porsche.*’

Relación: *neutral*. Realmente no podemos saber por la premisa, si lo que se afirma en la hipótesis sea verdadero o falso, ya que podría tener un Porsche o un Fiat.

NLI, anteriormente conocido como *Recognizing Textual Entailment* (Kovarchev y Taulé, 2022), es una de las tareas centrales del *Natural Language Understanding* (NLU), que es una rama del procesamiento del lenguaje natural. El NLU se suele definir mediante

⁸ Ejemplo tomado del artículo de 2022 de Kovarchev y Taulé

las tareas a las que se dedica, que tienen que ver con la forma en la que un sistema computacional entiende el lenguaje natural. Podemos encontrar dos tipos de tareas: las aplicaciones basadas en texto o basadas en diálogo (Allen, 1994). Esto implica que los sistemas deban interpretar diferentes aspectos lingüísticos como el análisis sintáctico y semántico, la resolución de la ambigüedad, la inferencia de significado o la interpretación de la intención del hablante entre otras.

Debido a la falta de recursos preexistentes, nos vimos obligados a crear un *corpus* desde cero para llevar a cabo esta tarea. Los ejemplos incluidos son frases inspiradas en el propio artículo original (Liu, Alisa, et al. 2023) y tomando la tabla de clasificación de palabras polisémicas, homónimas y monosémicas elaborada por (López-Cortés, 2023). El corpus se organizó en una tabla de clasificación en la que una de las columnas se utiliza para catalogar el tipo de ambigüedad de la que se trata, otra columna reservada para lo que consideramos que sería la premisa (en la que encontramos la ambigüedad), junto a ella una columna con la hipótesis y finalmente tres columnas en las que describimos el tipo de relación que tienen estos dos enunciados anteriores⁹. Hemos marcado solamente dos relaciones posibles para cada par de enunciados imitando el artículo de Liu, Alisa, et al. (2023). En la siguiente tabla podemos ver un ejemplo:

Tipo	Premisa	Hipótesis	Gold_Entailment	Gold_Contradiction
Léxica	Miguel nos espera en un banco	Miguel estaba cansado, no quería andar más	Miguel está sentado	Miguel ha ido a una entidad bancaria

Tabla 1. Tabla en la que se muestra un ejemplo de una de las ambigüedades contenida en corpus creado

Hemos trabajado con tres tipos concretos de ambigüedad: ambigüedad léxica, –*Esta alianza costó más de lo que esperaba*– (alianza: ‘anillo’; alianza: ‘unión’), que es un fenómeno que «se produce cuando una única palabra lleva aparejada varios significados.» (López-Cortés, 2023, pág. 22). Ambigüedad sintáctica –*Eligió un coche rápido*– (*rápido* como un modificador de coche: ‘eligió una caravana rápida’; o *rápido* analizado como un complemento de modo del verbo *elegir*: ‘rápidamente eligió el

⁹ Denominamos Gold_(relación) a aquellas verbalizaciones que hemos establecido manualmente para poder comparar con los resultados proporcionados por el Modelo de Lenguaje.

coche’), basada en la ambigüedad estructural, es decir, que podemos encontrar diferentes adjunciones entre los constituyentes de una frase, tal como se observó en las del ejemplo anterior.

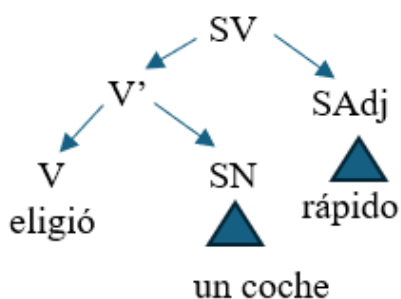


Figura 1. Árbol sintáctico con la 1ª interpretación

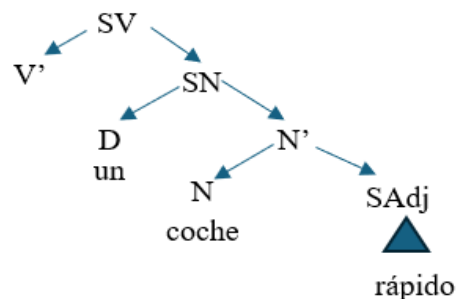


Figura 2. Árbol sintáctico con la 2ª interpretación

Por último, ambigüedad semántica basada en las diferencias de alcance de cuantificadores o de vaguedad referencial (definido vs. indefinido), entre otros fenómenos. Los cuantificadores, como señala la NGLE (RAE y ASELE, 2009) son palabras que «establecen cierta relación entre dos conjuntos». Los cuantificadores pueden expresar la medida de algo numéricamente (‘tres bolígrafos’) o nociones que no son reductibles a números, pero sí expresan cuantificación (‘pocos bolígrafos’). Por otra parte, cuando en una oración encontramos elementos pronominales denominamos a este fenómeno anáfora, –Mi amigo Juan **que** vivió en Zaragoza– (‘que’ es el elemento que hace referencia a Juan, mencionado previamente). En el caso en el que no sea claro a que elemento estamos haciendo referencia, lo denominamos de vaguedad referencial –Juan está hablando con María sobre la idea que tuvo– (‘que tuvo’ está siendo vago con respecto de quién es la idea). Así podemos incluir en la noción de ambigüedad semántica aquellos casos en los que encontramos «movimientos encubiertos de ascensión de cuantificadores» (Brucart y Gallego, 2012). Dentro de la propuesta de análisis por constituyentes del programa minimista de Chomsky, podemos encontrar dos tipos de movimientos: explícitos, es decir, que provocan una realización fonética diferente, o encubiertos. Movimiento encubierto es aquel que provoca un cambio en la interpretación de la oración, pero que no se realiza fonéticamente, por lo que no se puede hablar de un movimiento explícito como el que sucede en español y otras muchas lenguas con la ascensión del sintagma complementante (oraciones interrogativas) (Brucart y Gallego, 2012). Como señalan Brucart y Gallego (2012): «La justificación del movimiento encubierto es estrictamente semántica.» (2012: 38), y es por ello por lo que hemos optado por la

denominación de ambigüedad semántica para este tipo de fenómenos. A pesar de que podríamos aceptar que se trata de otro tipo de ambigüedad sintáctica debido a los movimientos (ocultos) de constituyentes expuestos anteriormente, nos resulta relevante la distinción entre tipos de oraciones para dedicarle una designación diferente.

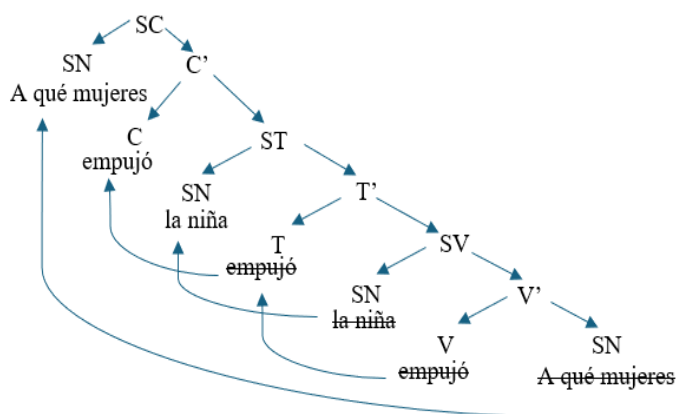


Figura 3. Ascensión con cambio fonológico

En la figura 3 podemos ver un ejemplo de una ascensión que sí provoca una realización fonética distinta a la que podríamos esperar en una oración no interrogativa. Se ha producido la ascensión del elemento marcado por el verbo como caso acusativo ('a qué mujeres'). Sin embargo, aunque los lingüistas postulan

movimientos sintácticos ocultos de los cuantificadores para explicar las dos interpretaciones posibles de expresiones como *Todos los alumnos admiran a un profesor* (todos al mismo o cada estudiante a uno distinto), parece más oportuno considerar estas ambigüedades como genuinamente semánticas.

En resumen, ciertas ambigüedades estructurales no se producen meramente por un posible doble análisis de constituyentes, sino que encontramos numerosas ambigüedades provocadas por las posibles interpretaciones semánticas que pueden tener los cuantificadores en una oración, como en uno de los ejemplos que examinaremos, del tipo de *Tres profesores examinan a cuatro estudiantes* (en el que una de las interpretaciones que podríamos encontrar sería que tres profesores examinen cada uno a los cuatro mismos alumnos o que tres profesores examinan a cuatro estudiantes diferentes)¹⁰.

También hemos recurrido a esta denominación para aquellas ambigüedades provocadas por una vaguedad referencial, y, por lo tanto, de carácter semántico –*Un estudiante votó a cada candidato*– (ya que 'un estudiante' puede referirse a uno específico o a uno cualquiera).

¹⁰ Este ejemplo, al igual que otros de esta categoría puede tener más interpretaciones.

Partimos, pues, de tres tipos de ambigüedades: la léxica (que se produce cuando una palabra tiene más de un sentido), la sintáctica (que se produce cuando hay más de un análisis sintáctico posible de una misma secuencia lineal y se basa en la sintaxis "visible" o "patente") y la semántica (que se produce cuando hay una diferencia en la estructura "no patente" o Forma Lógica de la sintaxis generativa). Nos parecía interesante esta distinción, ya que podría generar distintos retos al modelo de lenguaje utilizado y podríamos obtener diferentes conclusiones dependiendo de su rendimiento.

Nuestra intención al realizar este experimento no es conseguir una réplica exacta, si no la de tener un primer acercamiento con la elaboración de un corpus, conocer alguna de las metodologías en las que se pueden poner a prueba algunos modelos de lenguaje en el ámbito del PLN y obtener unos primeros resultados que podamos analizar. Hay que tener en cuenta que en el artículo de referencia la elaboración manual del corpus ha sido elaborada en dos fases: la primera, 141 ejemplos elaborados por 37 lingüistas y una segunda, llegando hasta 1645 ejemplos generados por un modelo de lenguaje. Mientras que en nuestro trabajo hemos realizado 19 ejemplos. Aunque también ha sido una gran forma de entrar en contacto con la anotación de NLI, dotándonos de una base sólida para comprender y aplicar esta tarea propia de PLN.

3. EXPERIMENTO

Tal y como se explica en el artículo de referencia, esta primera prueba consiste en estudiar si los modelos de lenguaje pueden aprender a generar desambiguaciones mediante un contexto. Este experimento se realiza mediante una técnica conocida como *prompting*. Es un procedimiento de aprendizaje supervisado que se utiliza en el campo del procesamiento del lenguaje natural que consiste en darle al modelo una orden escrita en lenguaje natural para que nos prediga la continuación de este (Lui, Pengfei, et al., 2023). Podemos encontrar diferentes tipos de *prompting*, pero nosotros vamos a utilizar la que se conoce como *chain-of-thought prompting* (Wei et al. 2022): se proporcionan ejemplos con razonamientos en la entrada del *prompt* para que el modelo siga el ejemplo.

Centrándonos en nuestro *prompt*, construimos un mensaje para el modelo en el que explicamos que hay una ambigüedad que hace que el entendimiento de una oración, a la que llamaremos *contexto*, es difícil de encontrar, pero que, mediante otra oración, a la que llamaremos *afirmación*, podemos llegar a resolver sin problemas. Para asegurarnos de que el modelo entiende la tarea, lo que hacemos es añadir ejemplos previamente resueltos,

creados en el corpus de datos. Lo que le vamos a pedir al modelo es que nos resuelva la ambigüedad dándonos una verbalización de la relación de NLI que entiende que tiene premisa e hipótesis.

Debemos destacar algunas diferencias entre el experimento original y el nuestro. En el artículo de referencia introducían 4 ejemplos en el *prompt*, mientras que nosotros hemos utilizado solo dos. Además, Liu et al. (2023) utilizan diferentes ejemplos para cada *prompt*. Hemos optado por emplear los mismos ejemplos para el mismo tipo de ambigüedad. Es decir, utilizaremos los ejemplos que se presentaron en las instrucciones de la tarea de resolución de ambigüedad léxica. Estos ejemplos se centrarán en resolver ambigüedades léxicas. No podemos estar seguros de que en el artículo original también dividan por tipo de ambigüedad por la imposibilidad de hacer un escrutinio intensivo de los ejemplos. Otra diferencia que podemos encontrar solamente en el *prompt* es que se le piden al modelo dos o tres interpretaciones para cada afirmación, nosotros le reducimos esa posibilidad a dos solamente.

Para la resolución de cada ejemplo se abría un diálogo nuevo en ChatGPT para evitar que haya una influencia de los datos y respuestas anteriores. Igualmente, también afecta la fecha en la que se realiza el experimento, ya que este modelo está en constante aprendizaje, por lo que los resultados pueden cambiar de un día para otro. Esta diferencia de resultados dependiendo del día en el que se realice la comentaremos más adelante.

Aquí dejamos una muestra en qué consiste nuestro *prompt*. La instrucción ha sido traducida directamente del inglés (con los mínimos cambios mencionados previamente), mientras que los ejemplos son los creados específicamente en nuestro *dataset* con el formato NLI:

Instrucción: Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Ejemplo:

Contexto: Miguel nos espera en un **banco**

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es **verdadera**.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es **falsa**.

Contexto: Esta **alianza** costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es **inconclusa**.

2- El narrador se va a casar. Por lo que la afirmación es **verdadera**.

Contexto: <PREMISA>

Afirmación: <HIPÓTESIS>

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Donde <PREMISA> e <HIPÓTESIS> serán sustituidos por la oración correspondiente con la columna de la tabla de corpus con este nombre.

Este ejemplo es el que hemos utilizado para aquellas oraciones que hemos identificado como ambigüedad léxica, ya que ambos ejemplos resuelven una ambigüedad de este tipo¹¹. Lo que esperamos que haga el modelo de lenguaje, en este caso, ChatGPT, es que nos dé dos oraciones en las que haga la misma tarea que le hemos mostrado en los ejemplos. Es decir, que si le hubiéramos pedido que resolviera el siguiente par de premisa-hipótesis (contexto-afirmación): **Contexto:** Perdió el **billete**. - **Afirmación:** Tuvo que comprar otro en la estación, lo que estaríamos esperando sería:

1- Perdió un billete de tren. Por lo que la información es **verdadera**.

2- Perdió dinero. Por lo que la premisa sería **falsa**.

Adelantándonos a posibles críticas que podemos encontrar en este tipo de prueba, vamos a comentar algunos aspectos que nos ha parecido que podían debilitar esta forma de evaluar. En primer lugar, da la impresión de que en la propia instrucción se prepara al modelo para la tarea de la que se va a encargar, es decir, se le pone en sobre aviso de que va a tener que resolver una ambigüedad que va a encontrar siempre en la frase denominada como contexto. Esto podría provocar que realmente no comprobemos la capacidad de resolución de ambigüedades, sino la capacidad creativa que puede tener el

¹¹ El resto de *prompts* se mostrarán en los siguientes apartados.

modelo de lenguaje a la hora de hacer una tarea. Además, al crear manualmente el *data set* que va a ser utilizado para poner a prueba el modelo, podemos crear pares de premisas e hipótesis que provoquen un sesgo en el modelo. Es decir, que no se puede estar seguro de si el modelo está verdaderamente desambiguando o si realmente está tomando la información que se le ha dado en la premisa, como los propios autores del artículo ponen de ejemplo¹² (Liu, Alisa et al., 2023):

P: He always ignores his mother's advice to follow his own dreams.

H: He follows his dreams.

Lo que va a hacer el modelo es tomar la primera afirmación (contexto o premisa) –Él ignora el consejo de su madre– y poner la etiqueta de falso o hacer una negación de la afirmación –Él no ignora el consejo de su madre– y poner la etiqueta de verdadero. Sin realmente resolver la ambigüedad que se le presenta: ‘Él ignora los consejos de su madre para perseguir sus sueños’ o ‘Él ignora los consejos de su madre de perseguir sus sueños’. «Por lo tanto, la corrección juzgada por los humanos puede sobreestimar la capacidad de los modelos para informar con precisión de la fuente de ambigüedad»¹³ (Liu, Alisa, et al., 2023).

También habría que tener en cuenta, si este tipo de ejemplos son fácilmente reconocibles por cualquier hablante o si realmente estamos intentando forzar al modelo de lenguaje a que haga algo que para los propios humanos supone también un reto. Una idea futura que se podría realizar como continuación de este trabajo sería la comparación del rendimiento de hablantes de español de distintos niveles y edades frente los resultados de ChatGPT, e incluso compararlo con los resultados que nos pueden ofrecer otros modelos de lenguaje.

4. RESULTADOS

En esta sección se van a presentar los resultados que hemos obtenidos al realizar el experimento. Están divididos según los tipos de ambigüedad que hemos especificado previamente. Además de mostrar las respuestas proporcionadas por el modelo y de centrarnos en la comprobación de correcto o incorrecto, hemos intentado buscar

¹² Este ejemplo solo funciona en inglés, ya que si lo traducimos podemos encontrar dos formas diferentes de expresarlo, reforzando la idea de que podemos encontrar una ambigüedad estructural en la oración.

¹³ «Thus, the human-judged correctness may overestimate the models' ability to precisely report the source of ambiguity.»

explicaciones a sus respuestas para intentar dilucidar de alguna manera la capacidad de desambiguación que puede llegar a tener ChatGPT.

4.1. AMBIGÜEDAD LÉXICA

En primer lugar, vamos a mostrar las respuestas que hemos recibido para la prueba de ambigüedad léxica. Las oraciones que hemos querido que resolviera son las siguientes:

Contexto: Hay que acabar con todos los **brotos**.

Afirmación: Hay una epidemia de una enfermedad

Contexto: Encontró un montón de **cartas** encima de la mesa.

Afirmación: Las cartas le llegaron por correo.

Contexto: Perdió el **billete**.

Afirmación: Tuvo que comprar otro en la estación.

Contexto: Dejó la **caña** en la mesa.

Afirmación: Se fue a recoger el resto de bebidas de la barra.

Nos hemos encontrado que, en este tipo de ambigüedad, generalmente, el modelo no encuentra apenas problemas a la hora de realizar la desambiguación. Esto lo podemos comprobar mediante los siguientes ejemplos:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es **verdadera**.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es **falsa**.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es **inconclusa**.

2- El narrador se va a casar. Por lo que la afirmación es **verdadera**.

Contexto: Dejó la caña en la mesa.

Afirmación: Se fue a recoger el resto de bebidas de la barra. Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

1- El contexto es una reunión en un bar. Por lo tanto, la afirmación es **verdadera**, ya que es probable que haya ido a recoger más bebidas.

2- El contexto es una pesca en un río. Por lo tanto, la afirmación es **falsa**, ya que no tiene sentido ir a recoger más bebidas en este contexto.

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es **verdadera**.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es **falsa**.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es **inconclusa**.

2- El narrador se va a casar. Por lo que la afirmación es **verdadera**.

Contexto: Hay que acabar con todos los brotes.

Afirmación: Hay una epidemia de una enfermedad

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

1- El billete es de transporte público. Por lo que la afirmación es **verdadera**.

2- El billete es de lotería. Por lo que la afirmación es **falsa**.

Aunque nos hemos encontrado con algunos casos en los que parece que ha intentado resolver una ambigüedad que tiene que ver con el mundo real, más que con la propia ambigüedad que le estamos presentando. Mostramos el ejemplo:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es **verdadera**.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es **falsa**.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es **inconclusa**.

2- El narrador se va a casar. Por lo que la afirmación es **verdadera**.

Contexto: Encontró un montón de **cartas** encima de la mesa.

Afirmación: Las cartas le llegaron por correo.

¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

1- Las cartas están en un lugar visible y accesible. Por lo que la afirmación es **verdadera**.

2- Las cartas podrían haber sido dejadas allí por alguien más, no necesariamente recibidas por correo. Por lo que la afirmación es **falsa**.

Quizás este es uno de los casos en los que los ejemplos escogidos no estarían lo suficientemente bien expresados como para que el modelo realice las inferencias que nosotros podemos realizar por nuestro conocimiento del mundo.

También podemos encontrar algunos ejemplos en los que simplemente falla por completo, como el siguiente:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es **verdadera**.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es **falsa**.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es **inconclusa**.

2- El narrador se va a casar. Por lo que la afirmación es **verdadera**.

Contexto: Hay que acabar con todos los **brotos**.

Afirmación: Hay una epidemia de una enfermedad.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

- 1- Se está hablando de jardinería y la frase se refiere a eliminar brotes no deseados de plantas. Por lo que la afirmación es **verdadera**.
- 2- Se está hablando de controlar la propagación de una enfermedad y la frase se refiere a brotes de contagio. Por lo que la afirmación es **inconclusa**.

En definitiva, en estos ejemplos hemos encontrado una coincidencia de 2 desambiguaciones correctas, con respecto a lo esperado. Hemos comparado los cuatro ejemplos con las respuestas creadas previamente (en la columna de GOLD_*relación), y aunque las respuestas no son exactamente iguales, son respuestas que consideramos válidas.

4.2. AMBIGÜEDAD SINTÁCTICA

Las oraciones de ambigüedad sintáctica son las siguientes:

Contexto: Es una radio de música popular.

Afirmación: Ponen artistas poco conocidos.

Contexto: El pavo está listo para comer.

Afirmación: El pavo está hambriento.

Contexto: La profesora de historia rusa.

Afirmación: Es experta en la época zarista.

Contexto: María vio a Julia con el telescopio.

Afirmación: A Julia le interesa mucho la astronomía.

En estos casos da la sensación de que el modelo ha encontrado otro tipo de ambigüedades para resolver que no tienen tanto que ver con la ambigüedad provocada por la estructura, sino por el significado de algunas palabras. Es decir, podríamos decir que ha intentado resolver este tipo de ambigüedad como si fueran ejemplos de ambigüedad léxica, a pesar de que los ejemplos que hemos introducido en el *prompt* son claramente resoluciones de ambigüedad estructural. Vamos a demostrarlo con los siguientes ejemplos:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame

2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China.

Afirmación: Los juguetes son de China.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El vendedor es de China. Por lo que la afirmación es **inconclusa**
- 2- Los juguetes se han fabricado en China. Por lo que la afirmación es **verdadera**.

Contexto: Eligió un coche rápido.

Afirmación: No pensó mucho en qué coche comprar.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El coche es veloz. Por lo que la afirmación es **inconclusa**.
- 2- Fue rápido eligiendo el coche. Por lo que la afirmación es **verdadera**.

Contexto: Es una radio de música popular.

Afirmación: Ponen artistas poco conocidos.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

El contexto puede ser interpretado de varias maneras:

La radio se enfoca en música popular, lo que podría implicar que ponen artistas poco conocidos dentro del género de la música popular. Por lo tanto, la afirmación sería **verdadera**.

La radio es popular pero también conocida por promover artistas emergentes y poco conocidos, independientemente del género. En este caso, la afirmación sería **falsa**, ya que están poniendo artistas poco conocidos, pero no necesariamente dentro del género de la música popular.

Como hemos podido comprobar en este caso ha realizado una desambiguación de la palabra ‘popular’, en lugar de la interpretación de ‘radio popular’, la radio es muy conocida, o ‘música popular’, la música que se escucha en esa radio es muy conocida.

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China.

Afirmación: Los juguetes son de China.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El vendedor es de China. Por lo que la afirmación es **inconclusa**
- 2- Los juguetes se han fabricado en China. Por lo que la afirmación es **verdadera**.

Contexto: Eligió un coche rápido.

Afirmación: No pensó mucho en qué coche comprar.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El coche es veloz. Por lo que la afirmación es **inconclusa**.
- 2- Fue rápido eligiendo el coche. Por lo que la afirmación es **verdadera**.

Contexto: El pavo está listo para comer.

Afirmación: El pavo está hambriento.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

1- El pavo está cocido y preparado para ser consumido. Por lo tanto, la afirmación es **falsa**, ya que los pavos no tienen la capacidad de experimentar hambre como lo haría un animal vivo.

2- “Listo para comer” podría implicar que el pavo está alimentado y listo para ser sacrificado, lo que podría llevar a la interpretación de que está hambriento antes de ser sacrificado. En este caso, la afirmación sería **verdadera**.

Este ha sido uno de aquellos casos que nos han generado numerosas dudas, ya que como podemos ver hay una parte del razonamiento que parece que ha sido correctamente desambiguada, pero también encontramos una inferencia innecesaria y que no ha sido nombrada en el contexto. Esto nos da la sensación de que el modelo está aplicando el

conocimiento previo a estas situaciones y no, realmente, haciendo un análisis de estas dos oraciones únicamente.

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China.

Afirmación: Los juguetes son de China.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El vendedor es de China. Por lo que la afirmación es **inconclusa**
- 2- Los juguetes se han fabricado en China. Por lo que la afirmación es **verdadera**.

Contexto: Eligió un coche rápido.

Afirmación: No pensó mucho en qué coche comprar.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El coche es veloz. Por lo que la afirmación es **inconclusa**.
- 2- Fue rápido eligiendo el coche. Por lo que la afirmación es **verdadera**.

Contexto: María vio a Julia con el telescopio.

Afirmación: A Julia le interesa mucho la astronomía.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

1- María vio a Julia utilizando el telescopio para observar el cielo nocturno. Por lo tanto, podría indicar un interés en la astronomía por parte de Julia. La afirmación es **verdadera**.

2- María vio a Julia con el telescopio, pero Julia podría haber estado utilizando el telescopio por otras razones, como mirar aves o edificios distantes. Por lo tanto, no podemos concluir con certeza que a Julia le interese mucho la astronomía. La afirmación es **inconclusa**.

En este ejemplo podemos ver como el modelo ha hecho inferencias a partir de una única estructura sintáctica, pero no resuelve la tarea que habíamos propuesto.

A pesar de esto, hemos podido comprobar que en algunos ejemplos es capaz de hacer una interpretación correcta:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China.

Afirmación: Los juguetes son de China.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El vendedor es de China. Por lo que la afirmación es **inconclusa**
- 2- Los juguetes se han fabricado en China. Por lo que la afirmación es **verdadera**.

Contexto: Eligió un coche rápido.

Afirmación: No pensó mucho en qué coche comprar.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El coche es veloz. Por lo que la afirmación es **inconclusa**.
- 2- Fue rápido eligiendo el coche. Por lo que la afirmación es **verdadera**.

Contexto: La profesora de historia rusa.

Afirmación: Es experta en la época zarista.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT:

La profesora enseña historia de Rusia. Por lo que la afirmación es **verdadera**.

La profesora es de origen ruso. Por lo que la afirmación es **inconclusa**.

Hemos considerado que, al ser un ejemplo muy común en manuales de lingüística, podría haber sido capaz de desambiguarlo correctamente debido a que ya conocía la respuesta de antemano.

4.3. AMBIGÜEDAD SEMÁNTICA

Las oraciones de ambigüedad semántica son las siguientes:

Contexto: Juan y Ana resolvieron el puzzle.

Afirmación: No resolvieron el puzzle juntos.

Contexto: Todos los estudiantes admiran a su madre.

Afirmación: La madre de uno de los estudiantes es una gran escritora.

Contexto: Tres profesores examinan a cuatro estudiantes.

Afirmación: Tres profesores examinan a cuatro estudiantes diferentes.

Contexto: Se cortó el pelo.

Afirmación: No fue a una peluquería.

Contexto: Un estudiante votó a cada candidato.

Afirmación: Todos los candidatos recibieron al menos un voto.

A pesar de nuestra primera intuición de que esta era la ambigüedad que más problemas podría generarle al modelo, ha resultado ser una de las ambigüedades en las que vemos una posible comprensión de los múltiples significados de las oraciones. En los siguientes ejemplos, podemos observar que, aunque no se resuelve la ambigüedad de la manera que podríamos esperar, es importante valorar en cada caso qué se está resolviendo realmente. Esto nos sugiere una cierta capacidad para realizar inferencias, aunque, nuevamente, no sean las que un hablante de español esperaría. Vamos a ver algunos ejemplos:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados.

Afirmación: Juan y Ana no son pareja.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Están casados el uno con el otro. Por lo que la afirmación es **falsa**.

2- Ambos están casados. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es **inconclusa**.

2- Los alumnos saben inglés y francés. Por lo que la afirmación es **verdadera**.

Contexto: Juan y Ana resolvieron el puzle.

Afirmación: No resolvieron el puzle juntos.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT

Juan y Ana resolvieron el puzle cada uno por su cuenta. Por lo que la afirmación es **verdadera**.

Juan y Ana colaboraron para resolver el puzle juntos. Por lo que la afirmación es **falsa**.

Encontramos algunos ejemplos en los que la respuesta es parcialmente correcta:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados.

Afirmación: Juan y Ana no son pareja.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Están casados el uno con el otro. Por lo que la afirmación es **falsa**.

2- Ambos están casados. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es **inconclusa**.
- 2- Los alumnos saben inglés y francés. Por lo que la afirmación es **verdadera**.

Contexto: Se cortó el pelo.

Afirmación: No fue a una peluquería.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT

Fue a una peluquería distinta a la habitual o se cortó el pelo en casa. Por lo que la afirmación es **verdadera**.

No se cortó el pelo en absoluto. Por lo que la afirmación es **falsa**.

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados.

Afirmación: Juan y Ana no son pareja.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Están casados el uno con el otro. Por lo que la afirmación es **falsa**.
- 2- Ambos están casados. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es **inconclusa**.
- 2- Los alumnos saben inglés y francés. Por lo que la afirmación es **verdadera**.

Contexto: Un estudiante votó a cada candidato.

Afirmación: Todos los candidatos recibieron al menos un voto.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT

Cada candidato recibió al menos un voto de algún estudiante. Por lo que la afirmación es **verdadera**.

Algunos candidatos podrían no haber recibido ningún voto, ya que no se especifica que cada candidato recibió al menos un voto. Por lo que la afirmación es **falsa**.

Este ejemplo podría ser problemático en su planteamiento, ya que, a diferencia de otros casos, no estamos resolviendo completamente la ambigüedad presente en el contexto. Sin embargo, esto refuerza la idea de que, aunque no detecta la ambigüedad, elabora una respuesta que aparenta ser correcta para cumplir con la tarea planteada.

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados.

Afirmación: Juan y Ana no son pareja.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Están casados el uno con el otro. Por lo que la afirmación es **falsa**.
- 2- Ambos están casados. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es **inconclusa**.
- 2- Los alumnos saben inglés y francés. Por lo que la afirmación es **verdadera**.

Contexto: Tres profesores examinan a cuatro estudiantes.

Afirmación: Tres profesores examinan a cuatro estudiantes diferentes.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT

Los tres profesores examinan a los mismos cuatro estudiantes. Por lo que la afirmación es **falsa**.

Los tres profesores examinan a diferentes grupos de cuatro estudiantes, pero no necesariamente a estudiantes diferentes en cada grupo. Por lo tanto, la afirmación es **verdadera**.

Y en otras que la respuesta es completamente incorrecta:

Prompt introducido:

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados.

Afirmación: Juan y Ana no son pareja.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Están casados el uno con el otro. Por lo que la afirmación es **falsa**.

2- Ambos están casados. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es **inconclusa**.

2- Los alumnos saben inglés y francés. Por lo que la afirmación es **verdadera**.

Contexto: Todos los estudiantes admiran a su madre.

Afirmación: La madre de uno de los estudiantes es una gran escritora.

Dado el contexto, ¿la afirmación es **verdadera**, **falsa** o **inconclusa**?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Respuesta de ChatGPT

Todos los estudiantes admiran a la madre de cada uno. Por lo que la afirmación es **verdadera**.

No todos los estudiantes necesariamente tienen una madre. Por lo que la afirmación es **inconclusa**.

En definitiva, no podemos negar o afirmar si ChatGPT tiene o no conocimiento sobre la ambigüedad. Aunque hemos podido comprobar mediante estos ejemplos que este modelo tiene problemas tanto con la sintaxis visible como con la invisible.

No hay que olvidar que este modelo, GPT-3.5, está en constante entrenamiento por el *feedback* y las conversaciones que mantiene con los usuarios. Esto quiere decir, que las respuestas a esta tarea pueden cambiar dependiendo del día en el que se hagan. Nos hemos quedado con las respuestas del día 23/05/2024, ya que son los resultados más recientes.

Tipo de ambigüedad	Correctas	Incorrectas	De los incorrectos, los que consideramos dudosos.	Nº de ejemplos
Léxica	2 (50%)	2 (50%)	0	4
Sintáctica	1 (25%)	3 (75%)	1 (33%)	4
Semántica	1 (20%)	4 (80%)	1 (25%)	5
Total	4 (30%)	9 (70%)	2 (22%)	13

Tabla 2. Se puede ver el número total de acierto (o error) para cada tipo de ambigüedad. Considerando también aquellos ejemplos que consideramos dudosos, pero que al no resolver la tarea marcamos como incorrectos.

5. CONCLUSIONES

En este trabajo hemos intentado comprobar cómo un modelo de lenguaje como ChatGPT realiza la desambiguación de distintas manifestaciones de oraciones ambiguas.

Como hemos mencionado previamente, nos hemos enfrentado a diversas limitaciones en el desarrollo del estudio. La principal es la falta de un corpus de datos que sea de tamaño comparable al original, lo cual es esencial para poder realizar comparaciones precisas y efectivas. Esta limitación específica, ya discutida previamente, surge de la necesidad de utilizar una anotación concreta, la propia de la disciplina de *Natural Language Inference*, que no ha sido lo suficientemente explorada ni desarrollada la lengua española. La falta de estos recursos y el adentrarse en un campo nuevo y desconocido ha

sido un reto adicional para la elaboración de un *data set* más preciso. Además, cabe destacar que el propio campo de estudio de las ambigüedades lingüísticas ha sido apartado en gran medida del estudio de las capacidades lingüísticas de los modelos de lenguaje. Esta exclusión se debe en gran parte a la complejidad inherente en el fenómeno lingüístico, ya que en numerosas ocasiones los hablantes de lenguas naturales no nos damos cuenta de que nos encontramos frente a ambigüedades constantemente, dado que no nos suponen ningún reto en la comprensión de los mensajes.

Pasando a la valoración de los resultados hemos podido observar cómo los modelos de lenguaje no tienen gran dificultad para realizar la desambiguación de palabras polisémicas a partir de una premisa y una hipótesis. En cambio, con las oraciones con ambigüedad sintáctica vemos que su rendimiento mejora cuando son frases que ha podido ver en su entrenamiento como “la profesora de historia rusa”, ya que es un ejemplo muy utilizado en manuales de sintaxis y, que podría estar presente en los corpus con los que se ha entrenado previamente a este modelo. En aquellas frases que le resultan novedosas, el modelo busca la ambigüedad en las propias palabras provocando casos como en ‘una radio de música popular’, que considere la palabra ‘popular’ como ambigua dando una respuesta que no esperábamos. En estos casos afirma que hay una ambigüedad porque lo hemos confirmado en el propio *prompt*, pero no parece que la encuentre y suple esto con la invención de una posible ambigüedad de otro tipo, que va más allá de lo que un humano suele hacer. Por último, en las que hemos denominado ambigüedades semánticas (por alcance de cuantificadores o vaguedad referencial) encontramos una capacidad de resolución de la tarea planteada más enigmática. Parece que el modelo tiene problemas tanto con la “sintaxis visible” como con la “invisible”. Produce una sensación de ser capaz de realizar una serie de inferencias, y aunque no haya evidencia de que asigne estructura sintáctica relevante para la ambigüedad, produce una sensación de conocimiento sobre el mundo, algo que podría deberse a la información que recibe en su entrenamiento.

A partir de estos resultados, surge una pregunta crucial: ¿es el modelo de lenguaje capaz de detectar la ambigüedad? Actualmente, no tenemos una respuesta definitiva, por lo que consideramos este experimento como un proyecto en desarrollo. Es necesario realizar más pruebas con diversos experimentos, intentando reducir los sesgos para determinar si los modelos de lenguaje pueden realmente resolver ambigüedades.

La creación de un corpus de datos estructurado en forma de *Natural language Inference* podría constituir una continuación de gran interés para este trabajo, dado el desafío que implica la búsqueda de ejemplos adecuados y el aprendizaje profundo en esta disciplina del Procesamiento de Lenguaje Natural. Asimismo, sería beneficioso evaluar un mayor número de modelos de lenguaje, ya que su rendimiento puede variar significativamente en función de cómo hayan sido entrenados. El artículo original, en el que nos basamos, realiza diferentes fases de prueba, las cuales podríamos replicar en futuros trabajos.

Debido a la extensión de este Trabajo de Fin de Grado y a nuestros conocimientos aún incipientes en el ámbito del PLN, no hemos podido llevar a cabo todos los experimentos propuestos en el artículo original. Sin embargo, esto no disminuye el interés del campo, ya que, como se mencionó anteriormente, la ambigüedad es un aspecto lingüístico que a menudo se ha relegado debido a su complejidad tanto para los modelos de lenguaje como para la búsqueda de métodos óptimos para evaluarlos.

Para concluir, este trabajo nos ha permitido profundizar en diversos aspectos relacionados con la creación de *corpus* y el *prompting* en el ámbito del PLN. A lo largo de nuestra investigación, hemos podido adquirir un conocimiento más sólido sobre estas técnicas y hemos aumentado nuestra inclinación por este campo, ya que el ámbito del Procesamiento del Lenguaje Natural está continuamente en desarrollo y resulta de gran interés seguir investigando y aprendiendo sobre él.

6. BIBLIOGRAFÍA

- ALLEN, James. *Natural Language Understanding*. Redwood City, CA: University of Rochester, 1994.
- BATES, Madeleine, Robert J. Bobrow y Ralph M. Weischedel. «Critical challenges for natural language processing.» Weischedel, Madeleine Bates y Ralph M. *Challenges in Natural Language Processing*. Cambridge: Cambridge University Press, 2006. 3-36.
- BRUCART, José M^a y Ángel J. Gallego. *El movimiento de constituyentes*. Madrid: Visor Libros, 2012.
- CHOWDHARY, K. R. «Natural Language Processing.» Chowdhary, K. R. *Fundamentals of Artificial Intelligence*. Jodhpur, India: Springer New Delhi, 2012. 603-649.
- ICARAZ VARÓ, Enrique y M.^a Antonia Martínez Linares. *Diccionario de lingüística moderna*. Barcelona: Ariel Lingüística, 2004.
- JURAFSKY, Daniel y James H. Martin. *Speech and Language Processing (3rd ed, draft)*. 2024.
- KOVATCHEV, Venelin y Mariona Taulé. «INFERS: A Natural Language Inference Corpus for Spanish Featuring Negation-Based Contrastive and Adversarial Examples.» *Proceedings of the 29th International Conference on Computational Linguistics*. 2022. 3873-3884.
- LIU, Alisa, y otros. «We're Afraid Language Models Aren't Modeling Ambiguity.» *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Ed. Association for Computational Linguistics. 2023. 790-807.
- LIU, Pengfei, y otros. «Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing.» *ACM Computing Surveys* 55.195 (2023): 1-35.
- LÓPEZ-CORTÉS, Natalia. *Homonimia y polisemia: la (des)conexión de los significados en español*. Salamanca: Ediciones Universidad Salamanca, 2023.
- MACCARTNEY, Bill. *Natural language inference. Ph. D. Dissertation*. Stanford, CA, USA: Stanford University, 2009.

- MIELKE, Sabrina J., et al. «Between words and characters: A brief history of open-vocabulary modeling and tokenization in NLP.» *arXiv preprint arXiv* (2021). <<https://arxiv.org/abs/2112.10508>>.
- OPENAI. «Openai Documentation.» s.f. *OpenAI*. 21 de 05 de 2024. <<https://platform.openai.com/docs/introduction>>.
- PIANTADOSI, Steven T., Harry Tily y Edward Gibson. «The communicative function of ambiguity in language.» *Cognition* 122.3 (2012): 280-291.
- PINKER, Steven y Paul Bloom. «Natural Language and Natural Selection.» *Behavioral and Brain Sciences*. Vol. 13. Cambridge University Press, 1990. 707-727.
- REAL ACADEMIA ESPAÑOLA, Asociación de Academias de la Lengua Española. *Nueva gramática de la lengua española*. Espasa Calpe España: Asociación de Academias de la Lengua Española, 2009-2011. Digital. <<https://www.rae.es/gram%c3%a1tica/>>.
- VASWANI, Ashish, et al. «Attention is all you need.» *Advances in neural information processing systems* 30 (2017).
- WEI, Jason, et al. «Chain-of-thought prompting elicits reasoning in large language models.» *Advances in neural information processing systems* 35 (2022): 24824-24837.

7. ANEXOS

CAPTURAS AMBIGÜEDAD LÉXICA



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- Miguel está sentado. Por lo que la afirmación es verdadera.
- 2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es falsa.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

- 1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es inconclusa.
- 2- El narrador se va a casar. Por lo que la afirmación es verdadera.

Contexto: Perdió el billete.

Afirmación: Tuvo que comprar otro en la estación

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1- El billete es de transporte público. Por lo que la afirmación es verdadera.

2- El billete es de lotería. Por lo que la afirmación es falsa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es verdadera.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es falsa.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es inconclusa.

2- El narrador se va a casar. Por lo que la afirmación es verdadera.

Contexto: Hay que acabar con todos los brotes.

Afirmación: Hay una epidemia de una enfermedad

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1- Se está hablando de jardinería y la frase se refiere a eliminar brotes no deseados de plantas. Por lo que la afirmación es verdadera.

2- Se está hablando de controlar la propagación de una enfermedad y la frase se refiere a brotes de contagio. Por lo que la afirmación es inconclusa.





Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es verdadera.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es falsa.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es inconclusa.

2- El narrador se va a casar. Por lo que la afirmación es verdadera.

Contexto: Encontró un montón de cartas encima de la mesa.

Afirmación: Las cartas le llegaron por correo.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa? No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1- Las cartas están en un lugar visible y accesible. Por lo que la afirmación es verdadera.

2- Las cartas podrían haber sido dejadas allí por alguien más, no necesariamente recibidas por correo. Por lo que la afirmación es falsa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Miguel nos espera en un banco

Afirmación: Miguel estaba cansado, no quería andar más

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- Miguel está sentado. Por lo que la afirmación es verdadera.

2- Miguel ha ido a una entidad bancaria. Por lo que la afirmación es falsa.

Contexto: Esta alianza costó más de lo que esperaba

Afirmación: El anillo ha costado mucho dinero

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1- El matrimonio ha sido por conveniencia. Por lo que la afirmación es inconclusa.

2- El narrador se va a casar. Por lo que la afirmación es verdadera.

Contexto: Dejó la caña en la mesa.

Afirmación: Se fue a recoger el resto de bebidas de la barra.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa? No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1- El contexto es una reunión en un bar. Por lo tanto, la afirmación es verdadera, ya que es probable que haya ido a recoger más bebidas.

2- El contexto es una pesca en un río. Por lo tanto, la afirmación es falsa, ya que no tiene sentido ir a recoger más bebidas en este contexto.



CAPTURAS AMBIGÜEDAD SINTÁCTICA



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China

Afirmación: Los juguetes son de China

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

El vendedor es de China. Por lo que la afirmación es inconclusa.

Los juguetes se han fabricado en China. Por lo que la afirmación es verdadera.

Contexto: Eligió un coche rápido

Afirmación: No pensó mucho en qué coche comprar

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

El coche es veloz. Por lo que la afirmación es inconclusa.

Fue rápido eligiendo el coche. Por lo que la afirmación es verdadera.

Contexto: Es una radio de música popular

Afirmación: Ponen artistas poco conocidos

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



El contexto puede ser interpretado de varias maneras:

La radio se enfoca en música popular, lo que podría implicar que ponen artistas poco conocidos dentro del género de la música popular. Por lo tanto, la afirmación sería verdadera.

La radio es popular pero también conocida por promover artistas emergentes y poco conocidos, independientemente del género. En este caso, la afirmación sería falsa, ya que están poniendo artistas poco conocidos pero no necesariamente dentro del género de la música popular.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China

Afirmación: Los juguetes son de China

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1. El vendedor es de China. Por lo que la afirmación es inconclusa.
2. Los juguetes se han fabricado en China. Por lo que la afirmación es verdadera.

Contexto: Eligió un coche rápido

Afirmación: No pensó mucho en qué coche comprar

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1. El coche es veloz. Por lo que la afirmación es inconclusa.
2. Fue rápido eligiendo el coche. Por lo que la afirmación es verdadera.

Contexto: El pavo está listo para comer.

Afirmación: El pavo está hambriento.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1. El pavo está cocido y preparado para ser consumido. Por lo tanto, la afirmación es falsa, ya que los pavos no tienen la capacidad de experimentar hambre como lo haría un animal vivo.
2. "Listo para comer" podría implicar que el pavo está alimentado y listo para ser sacrificado, lo que podría llevar a la interpretación de que está hambriento antes de ser sacrificado. En este caso, la afirmación sería verdadera.





Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China

Afirmación: Los juguetes son de China

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1. El vendedor es de China. Por lo que la afirmación es inconclusa.
2. Los juguetes se han fabricado en China. Por lo que la afirmación es verdadera.

Contexto: Eligió un coche rápido

Afirmación: No pensó mucho en qué coche comprar

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

1. El coche es veloz. Por lo que la afirmación es inconclusa.
2. Fue rápido eligiendo el coche. Por lo que la afirmación es verdadera.

Contexto: María vio a Julia con el telescopio.

Afirmación: A Julia le interesa mucho la astronomía

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



1. María vio a Julia utilizando el telescopio para observar el cielo nocturno. Por lo tanto, podría indicar un interés en la astronomía por parte de Julia. La afirmación es verdadera.
2. María vio a Julia con el telescopio, pero Julia podría haber estado utilizando el telescopio por otras razones, como mirar aves o edificios distantes. Por lo tanto, no podemos concluir con certeza que a Julia le interese mucho la astronomía. La afirmación es inconclusa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: El vendedor de juguetes de China

Afirmación: Los juguetes son de China

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

El vendedor es de China. Por lo que la afirmación es inconclusa.

Los juguetes se han fabricado en China. Por lo que la afirmación es verdadera.

Contexto: Eligió un coche rápido

Afirmación: No pensó mucho en qué coche comprar

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

El coche es veloz. Por lo que la afirmación es inconclusa.

Fue rápido eligiendo el coche. Por lo que la afirmación es verdadera.

Contexto: La profesora de historia rusa.

Afirmación: Es experta en época zarista.

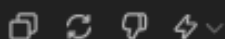
Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



La profesora enseña historia de Rusia. Por lo que la afirmación es verdadera.

La profesora es de origen ruso. Por lo que la afirmación es inconclusa.



CAPTURAS AMBIGÜEDADES SEMÁNTICAS

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados

Afirmación: Juan y Ana no son pareja

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Están casados el uno con el otro. Por lo que la afirmación es falsa.

Ambos están casados. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es inconclusa.

Los alumnos saben inglés y francés. Por lo que la afirmación es verdadera.

Contexto: Juan y Ana resolvieron el puzle.

Afirmación: No resolvieron el puzle juntos.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



Juan y Ana resolvieron el puzle cada uno por su cuenta. Por lo que la afirmación es verdadera.
Juan y Ana colaboraron para resolver el puzle juntos. Por lo que la afirmación es falsa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados

Afirmación: Juan y Ana no son pareja

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Están casados el uno con el otro. Por lo que la afirmación es falsa.

Ambos están casados. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es inconclusa.

Los alumnos saben inglés y francés. Por lo que la afirmación es verdadera.

Contexto: Se cortó el pelo.

Afirmación: No fue a la peluquería.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



Fue a una peluquería distinta a la habitual o se cortó el pelo en casa. Por lo que la afirmación es verdadera.

No se cortó el pelo en absoluto. Por lo que la afirmación es falsa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados

Afirmación: Juan y Ana no son pareja

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Están casados el uno con el otro. Por lo que la afirmación es falsa.

Ambos están casados. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es inconclusa.

Los alumnos saben inglés y francés. Por lo que la afirmación es verdadera.

Contexto: Un estudiante votó a cada candidato.

Afirmación: Todos los candidatos recibieron al menos un voto.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

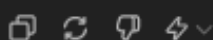
No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



El contexto puede ser interpretado de varias maneras:

Cada candidato recibió al menos un voto de algún estudiante. Por lo que la afirmación es verdadera.

Algunos candidatos podrían no haber recibido ningún voto, ya que no se especifica que cada candidato recibió al menos un voto. Por lo que la afirmación es falsa.



Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados

Afirmación: Juan y Ana no son pareja

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Están casados el uno con el otro. Por lo que la afirmación es falsa.

Ambos están casados. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es inconclusa.

Los alumnos saben inglés y francés. Por lo que la afirmación es verdadera.

Contexto: Tres profesores examinan a cuatro estudiantes.

Afirmación: Tres profesores examinan a cuatro estudiantes diferentes.

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



Los tres profesores examinan a los mismos cuatro estudiantes. Por lo tanto, la afirmación es falsa.

Los tres profesores examinan a diferentes grupos de cuatro estudiantes, pero no necesariamente a estudiantes diferentes en cada grupo. Por lo tanto, la afirmación es verdadera.

Para cada ejemplo, se te dará un contexto y una afirmación, en las que la correcta interpretación de la afirmación puede estar afectada por el contexto. Dame 2 interpretaciones del contexto que den lugar a diferentes interpretaciones sobre la afirmación.

Contexto: Juan y Ana están casados

Afirmación: Juan y Ana no son pareja

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Están casados el uno con el otro. Por lo que la afirmación es falsa.

Ambos están casados. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes de secundaria hablan dos lenguas.

Afirmación: Los alumnos hablan inglés y francés

Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:

Los alumnos hablan inglés, francés y otra lengua más. Por lo que la afirmación es inconclusa.

Los alumnos saben inglés y francés. Por lo que la afirmación es verdadera.

Contexto: Todos los estudiantes admiran a su madre.

Afirmación: La madre de uno de los estudiantes es una gran escritora.

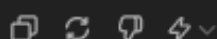
Dado el contexto, ¿la afirmación es verdadera, falsa o inconclusa?

No lo sabemos, porque el contexto puede ser interpretado de varias maneras:



Todos los estudiantes admiran a la madre de cada uno. Por lo que la afirmación es verdadera.

No todos los estudiantes necesariamente tienen una madre. Por lo que la afirmación es inconclusa.



TIPO	PREMISAS/CONTEXTO	HIPÓTESIS/AFIRMACIÓN	GOLD_INCONCLUSIVE	GOLD_TRUE	GOLD_FALSE	PREDICTED_INCONCLUSIVE	PREDICTED_TRUE	PREDICTED_FALSE
Léxica	Miguel nos espera en un banco	Miguel estaba cansado, no quería andar más		Miguel está sentado	Miguel ha ido a una entidad bancaria			
Léxica	Esta alianza costó más de lo que esperaba	El anillo ha costado mucho dinero	El matrimonio ha sido por conveniencia	El narrador se va a casar				
Léxica	Hay que acabar con todos los brotos .	Hay una epidemia de una enfermedad		Hay brotes de una enfermedad	Está creciendo mala hierba	Se está hablando de controlar la propagación de una enfermedad y la frase se refiere a brotes de contagio.	Se está hablando de jardinería y la frase se refiere a eliminar brotes no deseados de plantas.	
								Las cartas podrían haber sido dejadas allí por alguien más, no necesariamente recibidas por correo.
Léxica	Encontró un montón de cartas encima de la mesa	Las cartas le llegaron por correo	Las cartas son naipes	Las cartas son correspondencia			Las cartas están en un lugar visible y accesible.	El billete es de transporte público.
Léxica	Perdió el billete .	Tuvo que comprar otro en la estación		Perdió un billete de tren	Perdió dinero			
Léxica	Dejó la caña en la mesa	Se fue a recoger el resto de bebidas de la barra		Dejó su cerveza en la mesa	Dejó la caña de pescar encima de la mesa		El contexto es una reunión en un bar.	El contexto es una pesca en un río.
Sintáctica	El vendedor de juguetes de China	Los juguetes son de China	El vendedor es de China	Los juguetes se han fabricado en China				

Sintáctica	Eligió un coche rápido	No pensó mucho en qué coche comprar	El coche es veloz	Fue rápido eligiendo el coche				
							La radio se enfoca en música popular, lo que podría implicar que ponen artistas poco conocidos dentro del género de la música popular.	La radio es popular pero también conocida por promover artistas emergentes y poco conocidos, independientemente del género
Sintáctica	Es una radio de música popular	Ponen artistas poco conocidos		La radio es conocida por todos		La música es conocida por todos		
							"Listo para comer" podría implicar que el pavo está alimentado y listo para ser sacrificado, lo que podría llevar a la interpretación de que está hambriento antes de ser sacrificado.	El pavo está cocido y preparado para ser consumido.
Sintáctica	El pavo está listo para comer	El pavo está hambriento		El pavo quería comer		El pavo se podía comer		
	La profesora de historia rusa	Es experta en la época zarista	La profesora es rusa	Es profesora de historia de Rusia			La profesora enseña historia de Rusia	
Sintáctica	María vio a Julia con el telescopio	A Julia le interesa mucho la astronomía		Julia estaba con un telescopio	María estaba mirando	María vio a Julia con el telescopio, pero Julia podría haber	María vio a Julia utilizando el telescopio	

				por el telescopio	estado utilizando el telescopio por otras razones, como mirar aves o edificios distantes. Por lo tanto, no podemos concluir con certeza que a Julia le interese mucho la astronomía.	para observar el cielo nocturno. Por lo tanto, podría indicar un interés en la astronomía por parte de Julia.	
Semántica	Juan y Ana están casados	Juan y Ana no son pareja	Ambos están casados	Están casados el uno con el otro			
Semántica	Todos los estudiantes de secundaria hablan dos lenguas.	Todos los alumnos hablan inglés y francés	Los alumnos saben inglés y francés	Los alumnos hablan lenguas distintas			
Semántica	Juan y Ana resolvieron el puzle.	No resolvieron el puzle juntos	Ambos resolvieron el mismo puzle	Juan y Ana hicieron el puzle juntos		Juan y Ana resolvieron el puzle cada uno por su cuenta.	Juan y Ana colaboraron para resolver el puzle juntos.
Semántica	Todos los estudiantes admiran a su madre	La madre de uno de los estudiantes es una gran escritora	Los estudiantes admiran a la madre de ese estudiante		No todos los estudiantes necesariamente tienen una madre.	Todos los estudiantes admiran a la madre de cada uno.	
Semántica	Se cortó el pelo.	No fue a una peluquería	Se cortó el pelo a sí misma	Se cortó el pelo en la peluquería		Fue a una peluquería distinta a la habitual o se cortó el pelo en casa.	No se cortó el pelo en absoluto.

Semántica	Un estudiante votó a cada candidato	Todos los candidatos recibieron al menos un voto	Un mismo estudiante votó a todos los candidatos	Todos los candidatos fueron votados por al menos un estudiante			Cada candidato recibió al menos un voto de algún estudiante.	Algunos candidatos podrían no haber recibido ningún voto, ya que no se especifica que cada candidato recibió al menos un voto.
	Tres profesores examinan a cuatro estudiantes	Tres profesores examinan a cuatro estudiantes diferentes		Los mismos tres profesores examinan a los mismos cuatro estudiantes diferentes	Los mismos tres profesores examinan a los mismos cuatro estudiantes		Los tres profesores examinan a diferentes grupos de cuatro estudiantes, pero no necesariamente a estudiantes diferentes en cada grupo.	Los tres profesores examinan a los mismos cuatro estudiantes.