

UNIVERSIDAD DE COSTA RICA  
SISTEMA DE ESTUDIOS DE POSGRADO

APROXIMACIÓN LINGÜÍSTICA EN EL DISEÑO DE UN CORPUS ANOTADO EN  
ESPAÑOL SOBRE COVID-19 PARA SISTEMAS DE PREGUNTA-RESPUESTA

Tesis sometida a la consideración de la Comisión del Programa de Estudios de Posgrado de  
Lingüística para optar al grado y título de Maestría Académica en Lingüística

GRACIELA YISLÉN BARBOZA HIDALGO

Ciudad Universitaria Rodrigo Facio, Costa Rica

2023

## **DEDICATORIA**

Dedico esta tesis a mi esposo, Carl Albert Florin, y a mi hijo, Isaac Albert Florin Barboza. A ti, Carl, porque has sido mi apoyo incondicional en las buenas, en las malas y en las peores; porque has sido mi mano derecha, mi motivador personal y mi fuente de seguridad en momentos inciertos. A ti, mi amor, por estar aquí, por quedarte para saborear juntos otro logro académico y todos los que han de seguir llegando.

Te dedico este trabajo a ti, Isaac, pedacito de mí, porque has sido paciente, porque has sacrificado horas de juego sin mí, porque me has permitido demostrar lo que puedo ser capaz de hacer, sin importar cuánto tenga encima. Tú, mi pequeño, eres mi inspiración, mi luz y mi mayor hazaña. Eres la muestra perfecta de que yo puedo alcanzar imposibles cuando me lo propongo, de que no hay límites cuando trabajo por lo que quiero conseguir. Gracias a ti obtuve el mejor de los títulos: el de ser mamá.

## AGRADECIMIENTOS

En primer lugar, agradezco a Carl e Isaac por haberme dado el espacio para desarrollar mi investigación, por haber sido mi compañía, por haberme consolado en momentos de quebrantamiento y por haber compartido mis alegrías en los momentos de celebración. Hoy esta tesis es una realidad gracias a ustedes dos.

En segundo lugar, doy gracias al Dr. Jorge Antonio Leoni de León, por haber propuesto dirigir mi tesis de posgrado. La pandemia nos forzó a cambiar de tema, pero, gracias a ello, esta investigación va más acorde con nuestra pasión académica. Le agradezco haberme permitido aprender tanto a su lado con incontables horas en proyectos, traducciones e investigaciones juntos; estas experiencias le han dado forma a la lingüista que soy hoy. Gracias eternas.

En tercer lugar, agradezco a mis asesores, al Dr. Edgar Casasola Murillo y al Dr. Sergio Cordero Monge. Gracias, Edgar, porque su aporte computacional le dio sentido y rumbo a mi investigación. Gracias, Sergio, por ser un lector detenido, minucioso y comprometido con mi tema. Me honraron todos sus comentarios positivos y me enriquecieron sus correcciones. Gracias a los dos por su apoyo académico.

Agradezco a Dios, a mi familia y a mi amiga Karolina Sequeira Rodríguez, quienes me han acompañado en este proceso de formación, celebrando cada uno de mis logros y teniéndome paciencia cuando no he estado disponible social o emocionalmente. Gracias a mi suegra, Patricia Florin, por revisar mi resumen en inglés, por tan apropiadas correcciones, así como por sus nobles comentarios.

«Esta Tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Lingüística de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Lingüística»



---

Dra. Annia Espeleta Sibaja  
**Representante de la Decana**  
**Sistema de Estudios de Posgrado**



---

Dr. Jorge Antonio Leoni de León  
**Director de Tesis**



---

Dr. Edgar Casasola Murillo  
**Asesor**



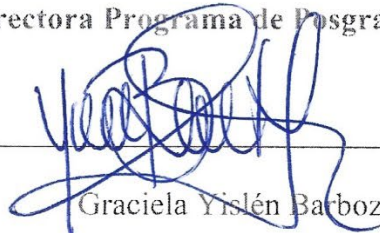
---

Dr. Sergio Cordero Monge  
**Asesor**



---

Dra. Gabriela Cruz Volio  
**Directora Programa de Posgrado en Lingüística**



---

Graciela Yislén Barboza Hidalgo  
**Sustentante**

## Tabla de contenidos

<b>DEDICATORIA .....</b>	<b>II</b>
<b>AGRADECIMIENTOS .....</b>	<b>III</b>
<b>RESUMEN .....</b>	<b>VII</b>
<b>ABSTRACT .....</b>	<b>VIII</b>
<b>LISTA DE TABLAS.....</b>	<b>IX</b>
<b>LISTA DE FIGURAS .....</b>	<b>X</b>
<b>LISTA DE ABREVIATURAS.....</b>	<b>XI</b>
<b>CAPÍTULO I. INTRODUCCIÓN .....</b>	<b>2</b>
1.1 <i>CONTEXTUALIZACIÓN DEL TEMA.....</i>	2
1.1.1 <i>Epidemia del COVID-19 .....</i>	6
1.2 <i>JUSTIFICACIÓN .....</i>	9
1.3 <i>PLANTEAMIENTO DEL PROBLEMA.....</i>	11
1.4 <i>OBJETIVOS DE LA TESIS.....</i>	12
1.5 <i>ALCANCES Y LIMITACIONES DE LA TESIS.....</i>	13
1.6 <i>CONTEXTO Y ANTECEDENTES A CORPUS LINGÜÍSTICOS.....</i>	15
1.6.1 <i>Procesamiento del lenguaje natural como método de extracción de la información.....</i>	15
1.6.2 <i>El enfoque simbólico .....</i>	17
1.6.3 <i>Corpus y Lingüística de Corpus.....</i>	18
1.6.4 <i>Antecedentes al corpus en español.....</i>	23
1.6.5 <i>Antecedentes del corpus en español con anotaciones semánticas .....</i>	26
1.6.6 <i>Modelos grandes de lenguaje .....</i>	27
1.6.7 <i>Estado de la cuestión .....</i>	31
1.7 <i>MARCO TEÓRICO.....</i>	48
1.7.1 <i>Las necesidades lingüísticas de los sistemas de pregunta-respuesta .....</i>	48
1.7.2 <i>Corpus y anotación lingüística .....</i>	49
1.7.3 <i>Roles temáticos .....</i>	54
1.7.4 <i>Relaciones sintáctico-semánticas de los argumentos verbales desde la Gramática Léxico- Funcional.....</i>	56
1.8 <i>METODOLOGÍA.....</i>	71
1.8.1 <i>El problema lingüístico.....</i>	72

1.8.2 Tipo de investigación.....	72
1.8.3 Descripción del conjunto de datos.....	74
1.8.4 Procesamiento del Corpus COVID-19.....	75
<b>CAPÍTULO II. DESARROLLO.....</b>	<b>87</b>
2.1 PRESENTACIÓN Y ANÁLISIS DE LOS RESULTADOS .....	87
2.1.1 Crear un subcorpus con los 200 verbos más frecuentes .....	87
2.1.2 Clasificar los argumentos verbales en cuanto a sus funciones temáticas y a sus funciones gramaticales.....	91
2.1.3 Desarrollar un modelo para el etiquetado de roles temáticos.....	106
2.1.4 Evaluación del acuerdo entre los anotadores.....	115
2.2 DISCUSIÓN DE LOS RESULTADOS .....	121
<b>CAPÍTULO III. CONCLUSIONES.....</b>	<b>124</b>
<b>BIBLIOGRAFÍA.....</b>	<b>127</b>
<b>ANEXO A .....</b>	<b>149</b>
<b>ANEXO B .....</b>	<b>150</b>
<b>ANEXO C.....</b>	<b>222</b>

## RESUMEN

Los recursos del Procesamiento del Lenguaje Natural, en conjunto con los corpus anotados, han acelerado el desarrollo de los sistemas inteligentes de pregunta-respuesta (*chatbots*), los cuales se entrenan para imitar el comportamiento lingüístico del ser humano. El etiquetado lingüístico de corpus es un proceso necesario en el entrenamiento de sistemas de pregunta-respuesta con métodos de aprendizaje automático; sin embargo, desde los inicios de la Inteligencia Artificial (AI) se ha intentado agilizar la evolución de la ingeniería lingüística a través de la automatización de las tareas pertenecientes al procesamiento del lenguaje natural, prescindiendo, muchas veces, del aporte de la Lingüística.

El etiquetado de roles semánticos en español es un tema que ha permanecido al margen en NLP y continúa con numerosos problemas sin resolver y, es por esto que, la intervención de lingüistas, con su conocimiento sobre la estructura interna de la lengua, permite mejorar y robustecer los modelos para *Machine Learning* con aportes teóricos lingüísticos pertinentes. Por lo anterior, en esta tesis se creó un modelo de anotación de roles temáticos en español, con un análisis descriptivo para los 200 verbos más frecuentes del *Corpus COVID-19*, empleando para esto la Lingüística de Corpus como metodología y la Gramática Léxico-Funcional (LFG) como base teórica.

Esta tesis se enfocó en el aspecto meramente lingüístico de la anotación, y no en la creación de *chatbots* ni en el entrenamiento puesta a punto (*fine-tuning*) para *ChatGPT*. Con este modelo de anotación se comparó el acuerdo entre los anotadores humanos y el de *ChatGPT*, por medio del coeficiente kappa de Fleiss. En este trabajo se concluye que *ChatGPT* obtuvo un desempeño inferior al de los humanos, con un valor de  $\kappa$  de 0.420 y con una precisión de 0.539, comparado al de los humanos que obtuvieron un valor de  $\kappa$  de 0.600, con una precisión de 0.700.

Esta investigación ha sido de carácter empírico, con pocos antecedentes para la lengua española, porque, al momento del desarrollo de la tesis, no se habían publicado trabajos en donde se compararan las anotaciones humanas y las anotaciones de un modelo de lenguaje de gran tamaño, como *ChatGPT*, para los roles temáticos; así como tampoco se encontraron antecedentes que ofrecieran una guía claramente replicable para anotar los roles temáticos. Su diseño fue exploratorio, por tratarse de una propuesta de anotación de etiquetas lingüísticas para roles temáticos en español, y el método de análisis fue por medio del análisis textual de corpus en español, empleando la perspectiva de la Lingüística de Corpus.

## ABSTRACT

Natural Language Processing (NLP) resources, together with annotated corpora, have accelerated the development of intelligent question-answer systems (chatbots), which are trained to imitate human linguistic behavior. Corpus linguistic annotation is a necessary process in training question-answer systems with machine learning methods. However, since the beginnings of Artificial Intelligence (AI), attempts have been made to speed up the evolution of linguistic engineering through the automation of tasks belonging to natural language processing, often disregarding the contribution of Linguistics.

The labeling of semantic roles in Spanish is a topic that has remained on the sidelines in NLP and continues with numerous unresolved problems. For this reason, the intervention of linguists, with their knowledge of the internal structure of the language, allows to improve and strengthen the models for Machine Learning with relevant linguistic theoretical contributions. Due to the above, in this thesis a thematic role annotation model in Spanish was created, with a descriptive analysis for the 200 most frequent verbs of the *Corpus COVID-19*, using for this the Corpus Linguistics as methodology and the Lexical-Functional Grammar (LFG) as the theoretical basis.

This thesis focused on the purely linguistic aspect of the annotation, and not on the creation of chatbots or fine-tuning training for *ChatGPT*. With this annotation model, the agreement of the human annotators and that of *ChatGPT* was compared using the Fleiss kappa coefficient. In this work it is concluded that *ChatGPT* obtained a lower performance than that of human annotators, with a  $\kappa$  value of 0.420, with a precision of 0.539; compared to that of humans who obtained a  $\kappa$  value of 0.600, with a precision of 0.700.

This research has been of an empirical nature, with few precedents for the Spanish language, because, at the time of the development of the thesis, no papers had been published comparing human annotations and those done by large language model, such as *ChatGPT*, for the thematic roles; nor were found antecedents that offered a clearly replicable guide to tag the thematic roles. Its design was empirical, since it is a proposal for the annotation of linguistic labels for thematic roles in Spanish, and the method of analysis was through the textual analysis of corpus in Spanish, using the perspective of Corpus Linguistics.

## LISTA DE TABLAS

Tabla 1. <i>Análisis lingüístico completo de FreeLing</i> .....	3
Tabla 2. <i>Catálogo de roles semánticos</i> .....	54
Tabla 3. <i>Descomposición de características de funciones de argumento</i> .....	68
Tabla 4. <i>Directriz para interpretar los valores</i> .....	85
Tabla 5. <i>Datos del Corpus C0095 tomados de Sketch Engine</i> .....	87
Tabla 6. <i>Muestra de la lista de frecuencias verbales en el Corpus COVID-19</i> .....	88
Tabla 7. <i>Registro de ID para archivos del subcorpus COVID-19</i> .....	90
Tabla 8. <i>Lista de frases por verbo descargada de Sketch Engine</i> .....	90
Tabla 9. <i>Categorías de la estructura-c</i> .....	96
Tabla 10. <i>Análisis en la estructura-c para oración subordinada</i> .....	96
Tabla 11. <i>Clasificación del sintagma verbal SV en la estructura-c</i> .....	96
Tabla 12. <i>Etiquetado de las oraciones subordinadas en la estructura-c</i> .....	97
Tabla 13. <i>Clasificación de la estructura-f</i> .....	98
Tabla 14. <i>Etiquetado de XCOMP en la estructura-f</i> .....	98
Tabla 15. <i>Etiquetado de XCOMP Prep en la estructura-f</i> .....	99
Tabla 16. <i>Casos de 'CAG' en el Corpus COVID-19</i> .....	99
Tabla 17. <i>Distribución de etiquetas para los argumentos primarios</i> .....	106
Tabla 18. <i>Distribución de etiquetas para los modificadores verbales</i> .....	106
Tabla 19. <i>Distribución de las etiquetas por rol temático</i> .....	107
Tabla 20. <i>Distribución de los ítemes según el anotador</i> .....	108
Tabla 21. <i>Distribución de los grupos de anotadores</i> .....	109
Tabla 22. <i>Respuestas de los anotadores humanos</i> .....	109
Tabla 23. <i>Respuestas de ChatGPT a la encuesta</i> .....	111
Tabla 24. <i>Kappa de Fleiss para los anotadores humanos</i> .....	115
Tabla 25. <i>Kappa de Fleiss para anotadores humanos junto con ChatGPT</i> .....	116
Tabla 26. <i>Acuerdo entre 1 anotador humano y ChatGPT</i> .....	118
Tabla 27. <i>Comparación del valor de <math>\kappa</math> al combinar ChatGPT con 2 anotadores humanos</i> .....	118
Tabla 28. <i>Comparación del valor de <math>\kappa</math> al combinar ChatGPT con 3 anotadores humanos</i> .....	119
Tabla 29. <i>Comparación del valor de <math>\kappa</math> al combinar ChatGPT con 4 anotadores humanos</i> .....	119
Tabla 30. <i>Comparación del valor de <math>\kappa</math> para los anotadores humanos y para ChatGPT</i> .....	120

## LISTA DE FIGURAS

Figura 1. <i>Ejemplo de Viber de WHO</i> .....	8
Figura 2. <i>Modelización del clustering jerárquico aglomerativo</i> .....	34
Figura 3. <i>Léxico verbal para el verbo VALORAR</i> .....	36
Figura 4. <i>Estructura de constituyentes</i> .....	37
Figura 5. <i>Estructura de dependencias</i> .....	37
Figura 6. <i>Estructura semántica</i> .....	38
Figura 7. <i>Representación sintáctica profunda y semántica</i> .....	40
Figura 8. <i>Árbol sintáctico de «El aro de la bicicleta rodaba sin problemas»</i> .....	52
Figura 9. <i>Formato de las reglas en LFG</i> .....	58
Figura 10. <i>Árbol de constituyentes anotado con esquemas funcionales</i> .....	59
Figura 11. <i>Asociación entre los nombres de las estructuras-c y sus nodos</i> .....	60
Figura 12. <i>Consolidación del algoritmo de instanciación</i> .....	61
Figura 13. <i>Representación de la estructura-f</i> .....	62
Figura 14. <i>Correspondencia entre la estructura-c y la estructura-f</i> .....	63
Figura 15. <i>Constructo léxico-sintáctico</i> .....	65
Figura 16. <i>Correspondencia entre las estructuras-c, -f y -a para Carl bebe café</i> .....	66
Figura 17. <i>Distintos tipos de funciones</i> .....	67
Figura 18. <i>Etiquetas de los roles temáticos de los argumentos del verbo</i> .....	78
Figura 19. <i>Etiquetas de los adjuntos o modificadores verbales</i> .....	78
Figura 20. <i>Fórmula del coeficiente kappa de Fleiss</i> .....	83
Figura 21. <i>Fórmula para calcular el acuerdo observado entre los anotadores</i> .....	84
Figura 22. <i>Fórmula para calcular el acuerdo esperado entre los anotadores</i> .....	84
Figura 23. <i>División de los URL de los documentos del Corpus C0095</i> .....	88
Figura 24. <i>Plataforma del Léxico verbal del corpus SenSem (GRIAL)</i> .....	92
Figura 25. <i>Inconsistencias en el Léxico verbal del corpus SenSem (GRIAL)</i> .....	93
Figura 26. <i>Inconsistencias en el etiquetado del GRIAL</i> .....	94
Figura 27. <i>Archivo para el verbo ACABAR con el análisis de roles en TextServer</i> .....	95
Figura 28. <i>Ejemplo de respuesta de ChatGPT-4 para la encuesta</i> .....	111
Figura 29. <i>Variación de las respuestas de ChatGPT-3.5 con el mismo prompt</i> .....	112
Figura 30. <i>Alucinaciones de ChatGPT</i> .....	113

## LISTA DE ABREVIATURAS

<b>AI</b>	Inteligencia Artificial ( <i>Machine Learning</i> )
<b>BERT</b>	Representaciones de codificador bidireccional de transformadores ( <i>Bidirectional Encoder Representations from Transformers</i> )
<b>DL</b>	<i>Deep Learning</i>
<b>GGT</b>	Gramática Generativa Transformacional ( <i>Transformational-generative Grammar</i> )
<b>GLUE</b>	Evaluación general de comprensión del lenguaje
<b>GPT</b>	Transformador preentrenado generativo ( <i>Generative Pretrained Transformer</i> )
<b>HMM</b>	Modelos ocultos de Markov ( <i>Hidden Markov Models</i> )
<b>HPSG</b>	Gramática Sintagmática Nuclear ( <i>Head-driven phrase structure grammar</i> )
<b>IAA</b>	Acuerdo entre anotadores ( <i>Inter Annotator Agreement</i> )
<b>LC</b>	Lingüística de Corpus
<b>LComp</b>	Lingüística Computacional
<b>LFG</b>	Gramática Léxico-Funcional ( <i>Lexical-Functional Grammar</i> )
<b>LLM</b>	Modelos de lenguaje de gran tamaño ( <i>Large Language Model</i> )
<b>ML</b>	Aprendizaje automático ( <i>Machine Learning</i> )
<b>NER</b>	Reconocimiento de Nombre de Entidades ( <i>Named Entity Recognition</i> )
<b>NLP</b>	Procesamiento del Lenguaje Natural ( <i>Natural Language Processing</i> )
<b>POS</b>	Etiquetado gramatical ( <i>Part-of-speech tagging</i> )
<b>SQuAD</b>	Conjunto de datos estándar de respuesta a preguntas ( <i>Standard Question Answering Dataset</i> )
<b>URL</b>	<i>Uniform Resource Locator</i>

# CAPÍTULO I. INTRODUCCIÓN

En este capítulo se formulan los aspectos de la presente tesis en diferentes apartados, a saber: la contextualización del tema en 1.1, la justificación en 1.2, el planteamiento del problema en 1.3, los objetivos en 1.4, los alcances y las limitaciones de este proyecto en 1.5 y el estado de la cuestión en 1.6. Asimismo, se explican el marco teórico en 1.7 y la metodología que se siguió para dar forma a este trabajo académico en 1.8.

## 1.1 Contextualización del tema

La Lingüística de Corpus (LC) es una metodología para tratar, objetiva y estadísticamente, la lengua en uso, a fin de estudiar cualquier dominio lingüístico presente en corpus compilados para tal efecto (Parodi, 2010). Con la LC se obtienen datos valiosos que sirven como evidencia científica para explicar fenómenos del lenguaje (Wallis y Nelson, 2001) y para descifrar lo que los hablantes hacen realmente con su lengua, porque, contrario a la tradición lingüística convencional teórica, el comportamiento del hablante no se deduce a partir de estándares hipotéticos del lenguaje (Culicover y Jackendoff, 2005), sino de su uso práctico. La LC es uno de los principales métodos que se emplean en la actualidad para entrenar sistemas computacionales en el procesamiento del lenguaje natural (NLP, por sus siglas en inglés) y en el Aprendizaje Automático (AA) o *Machine Learning* (ML) en inglés (Navarro-Colorado, 2007).

Los corpus son una colección representativa y balanceada de textos legibles por computadoras, que se han producido en un entorno de comunicación natural tomada directamente de los usuarios mismos de la lengua en estudio (Pustejovsky y Stubbs, 2013; Rojo, 2014, 2016, 2021; Völkel y Kretzschmar, 2021). Los corpus se crean con material seleccionado proveniente de textos diversos, los cuales pueden ser escritos, orales o visuales (Wallis y Nelson, 2001), dependiendo del interés investigativo, como el *Brown Corpus* (University of Essex, 1998), el *Penn TreeBank* (Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., Taylor, A., 1999), los creados por la Real Academia Española (*CREA*, *CORDE*, *CORPES*, etc.), el *SenSem* (Fernández y Vázquez, 2015) o el *Ancora* (Taulé, Martí y Recasens, 2008). En los antecedentes se ofrecen estos y otros casos de corpus lingüísticos como referencia.

La anotación de los textos de un corpus es una práctica común para describir una lengua y, en particular, para proporcionar algún tipo de explicación gramatical, sea esta sintáctica, semántica, morfológica, fonológica, fonética, léxica, discursiva, pragmática o de estructura textual,

dependiendo del objeto de estudio para el que se compiló el corpus (Wallis y Nelson, 2001; Pustejovsky y Stubbs, 2013). Las anotaciones pueden ser hechas por humanos, normalmente expertos en el tema que se busca profundizar o, por programas computacionales de manera automática (no supervisada) o semiautomática (supervisada). No obstante, los datos que se obtienen de estos sistemas no siempre son confiables (Newman y Cox, 2020; Spinde, 2021; Moeller y Hulden, 2021; Reiss, 2023).

A continuación, en la Tabla 1, se presenta como ejemplo el análisis completo (*full parsing*) de la frase en español «Por eso es muy importante estar atentos a las señales de nuestro cuerpo ante el posible contagio» (La Nación, 2021) conseguida de manera automatizada:

**Tabla 1**

*Análisis lingüístico completo de FreeLing*

▼ CONLL format

1	Por	por	SP	SP	- - -	(grup-verb:3(grup-sp:1(pre:1)	- - - -
2	eso	ese	PD00S00	PD	- - -	(sn:2(pron-ms:2(pdem-ms:2))))	- - - -
3	es	ser	VSIP3S0	VSI	- -	02604760-v (verb:3)	- - - -
4	muy	muy	RG	RG	- -	00014285-r (s-adj:5(sadv:4)	- - - -
5	importante	importante	AQ0CS00	AQ	- -	00624026-a (s-a-ms:5(a-ms:5)))	- - - -
6	estar	estar	VMN0000	VMN	- -	02604760-v (grup-verb-inf:6(infinitiu:6(forma-inf:6)))	- - - -
7	atentos	atento	AQ0MP00	AQ	- -	01194483-a (s-adj:7(s-a-mp:7(a-mp:7)))	- - - -
8	a	a	SP	SP	- - -	(grup-sp:8(pre:8)	- - - -
9	las	el	DA0FP0	DA	- - -	(sn:10(espec-fp:9(j-fp:9))	- - - -
10	señales	señal	NCFP000	NC	- -	06791372-n (grup-nom-fp:10(n-fp:10))	- - - -
11	de	de	SP	SP	- - -	(sp-de:11	- - - -
12	nuestro	nuestro	DP1MSP	DP	- - -	(sn:13(espec-ms:12(pos-ms:12))	- - - -
13	cuerpo	cuerpo	NCMS000	NC	- -	05216365-n (grup-nom-ms:13(n-ms:13))))))	- - - -
14	ante	ante	SP	SP	- - -	(grup-sp:14(pre:14)	- - - -
15	el	el	DA0MS0	DA	- - -	(sn:17(espec-ms:15(j-ms:15))	- - - -
16	posible	posible	AQ0CS00	AQ	- -	01821266-a (grup-nom-ms:17(s-a-ms:16(a-ms:16))	- - - -
17	contagio	contagio	NCMS000	NC	- -	07360647-n (grup-nom-ms:17(n-ms:17))))))	- - - -
18	.	.	Fp	Fp	- - -	(F-term:18)	- - - -

El etiquetado en la Tabla 1 se obtuvo del *parser* de *Freeling*<sup>1</sup> (Lexical Computing, 2020) y se expresa en formato CoNLL, el cual es un set de convenciones comunes establecido por el *Conference of Natural Language Learning* (Deng, 2023). Estas convenciones son generalmente seguidas por los desarrolladores de la tecnología del lenguaje; no obstante, estas convenciones no son estáticas, sino que varían con los años (Chiarcos, Ionov, Glaser y Fäth, 2021). Arriba, en la Tabla 1, cada palabra es analizada de acuerdo con su función en la frase y las etiquetas con sus códigos particulares son asignadas automáticamente por este *parser*. En el apartado del marco teórico se ampliará sobre este tema.

<sup>1</sup> Página accedida el 24 de abril de 2022: <https://nlp.lsi.upc.edu/freeling/demo/demo.php>

Navarro-Colorado (2007) plantea que la anotación lingüística es una forma de entender la lengua y que esta explicación de la lengua debe ser llevada a cabo por especialistas en lingüística. Muchos corpus son etiquetados para la construcción de diccionarios y tesauros, como CREA, CORDE o CORPES de la Real Academia Española, mientras que otros tantos han sido etiquetados con análisis lingüísticos de la estructura de cada frase por medio de árboles sintácticos gramaticales (Wallis y Nelson, 2001), como el *Corpus Ancora-Es* (Taulé, Martí y Recasens, 2008), el *IULA Treebank* (Marimon, Fisas, Bel, Vivaldi, Torner, Lorente, ... y Villegas, 2012) y el *LSP Spanish treebank*, desarrollados en el IULA de la Universidad Pompeu Fabra (Marimon, Bel, Fisas, Arias, Vázquez, Vivaldi, J., Morell y Lorente, 2014; Marimon y Bel, 2015). En los antecedentes se detallarán estos y otros ejemplos de corpus anotados en español.

Los corpus que han sido analizados (*parsed*) han creado un puente que vincula la teoría lingüística con los datos característicos del corpus en estudio y, este puente, a su vez, ha sido potenciado con algoritmos computacionales que perfeccionan la capacidad del análisis para confirmar o descartar afirmaciones teóricas y para buscar resultados significativos usando técnicas de ML, como sugieren Wallis y Nelson (2001).

Para que un sistema de ML tenga un desempeño de alta calidad es necesario que los corpus empleados hayan sido anotados y supervisados apropiadamente por expertos y no únicamente por métodos automatizados (Newman y Cox, 2020), porque los lingüistas tendrán la debida formación en las interpretaciones léxica, sintáctica, semántica, discursiva y pragmática de la lengua (Periñán Pascual, 2012). Es por esta razón que las contribuciones fundamentadas en teorías gramaticales que explican y sustentan dicha interpretación juegan un papel indispensable en el procesamiento del lenguaje natural y en el aprendizaje automático.

Los recursos del NLP, en conjunto con los corpus anotados, han acelerado el desarrollo de los sistemas inteligentes de pregunta-respuesta, llamados también *chatbots*. Estos son programas computacionales que se entrenan para imitar el comportamiento lingüístico del ser humano, es decir, para comunicarse con otros humanos por medio del lenguaje natural. Los orígenes de los *chatbots* se remontan a 1950, cuando Alan Turing propuso el *Turing test* o *Imitation Game* (Turing, 1950, 1956), con el cual se comprobaba si una computadora podía ‘pensar’ por sí misma y, principalmente, si esta podía mostrar rasgos humanos que fueran aceptados como válidos por otros humanos (Hodges, s.f.). A partir de la propuesta de Turing, muchos otros investigadores han intentado probar esta hipótesis, como Joseph Weizenbaum con ELIZA: «un procesador de

traducción en el sentido técnico de programación» (Weizenbaum, 1966, p. 43), que era capaz de mantener el hilo de la conversación, pero incapaz de sacar conclusiones válidas de lo que se le decía por su propia cuenta.

El entrenamiento de los *chatbots* se lleva a cabo por medio de dos métodos: los basados en reglas y los basados en Inteligencia Artificial (AI por sus siglas en inglés) (Navarro-Colorado, 2007; Pustejovsky y Stubbs, 2013; Coniam, 2014; Shum, He y Li, 2018; Veglis y Maniou, 2019; Gupta, Hathwar y Vijayakumar, 2020). Los del primer método se caracterizan por responder a comandos específicos y por ser incapaces de comprender el lenguaje natural, como es el caso del ayudante virtual del Banco de Costa Rica<sup>2</sup> que responde únicamente de acuerdo con las preguntas que han sido previamente establecidas por el programador y a términos propios de la jerga bancaria que han sido determinados como su léxico.

Existen diversos tipos de *chatbots*, dependiendo de la razón para la cual hayan sido creados, como los de texto breve con funciones variadas, desde informacionales, transaccionales, para control de dispositivos o para productividad empresarial (Veglis y Maniou, 2019); o aquellos para conversaciones largas, como los del segundo método basados en AI, que aprenden de las conversaciones que establecen con las personas y continuamente se actualizan sobre el uso de la lengua (Veglis y Maniou, 2019; Völkel y Kretzschmar, 2021), por lo que son capaces de generar mejores y más convincentes respuestas para los humanos.

Ejemplos de *chatbots* desarrollados con AI son *Siri*<sup>3</sup>, *Alexa*<sup>4</sup>, *Cortana*<sup>5</sup>, el *Asistente de Google*<sup>6</sup>, *LaMDA*<sup>7</sup>, *BERT (Bidirectional Encoder Representations from Transformers)* (Devlin y Chang, 2018; Devlin, Chang, Lee y Toutanova, 2019), *ChatGPT*<sup>8</sup>, *Microsoft Bing ChatGPT*<sup>9</sup>,

---

<sup>2</sup> Disponible en <https://bcrchat.bancobcr.com/BCR/chat.aspx>

<sup>3</sup> Disponible en <https://www.apple.com/es/siri/>. [Última consulta el 30 de marzo de 2023].

<sup>4</sup> Disponible en <https://developer.amazon.com/es-ES/alexa>. [Última consulta el 30 de marzo de 2023].

<sup>5</sup> Disponible en <https://support.microsoft.com/es-es/topic/-qu%C3%A9-es-cortana-953e648d-5668-e017-1341-7f26f7d0f825>. [Última consulta el 30 de marzo de 2023].

<sup>6</sup> Disponible en [https://assistant.google.com/intl/es\\_es/](https://assistant.google.com/intl/es_es/). [Última consulta el 30 de marzo de 2023].

<sup>7</sup> Cabe mencionar que *LaMDA* ha estado causando muchos comentarios en el mundo científico en estos días, desde que se han dado reportes de que tiene conciencia propia; aunque esta afirmación ha sido desmentida por expertos (Sparkes, 2022). Disponible en <https://blog.google/technology/ai/lamda/>. [Última consulta 30 de marzo de 2023].

<sup>8</sup> Disponible en <https://openai.com/blog/chatgpt>. [Última consulta el 28 de marzo de 2023].

<sup>9</sup> Disponible en <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>. [Última consulta el 28 de marzo de 2023].

*Jasper*<sup>10</sup>, *YouChat*<sup>11</sup> o *Chatsonic*<sup>12</sup>, entre otros que se han ido popularizando. Existen otros programas con quienes también se pueden mantener conversaciones más abiertas, más “humanas” y sobre casi cualquier temática, como *XiaoIce*<sup>13</sup>, un *chatbot* social que tiende a ser más atractivo para algunos, porque provee una conexión social más semejante a la humana (Zhou, Gao, Li y Shum, 2020). Otros sistemas de pregunta-respuesta menos conocidos, pero no necesariamente menores en desempeño son *Rytr*<sup>14</sup>, *INK*<sup>15</sup>, *Perplexity*<sup>16</sup> o *Google Bard*<sup>17</sup>,

El desempeño de *ChatGPT*, en particular, ha causado revuelo entre académicos, profesionales y el público lego desde su debut en 2022. Desde entonces este *chatbot* se ha venido poniendo a prueba desde diversos ángulos, con el fin de determinar su valor en la automatización en varias áreas, como en los negocios (George, George y Martin, 2023), en la academia (Lund y Wang, 2023; Chomsky, Roberts y Watumull, 2023, 8 de marzo; Johnson, 2023) y, por supuesto, en NLP (Oh y Shuler, 2022; Hariri, 2023; Kocoń, Cichecki, Kaszyca, Kochanek, Szydło, Baran, ... y Kazienko, 2023; Koubaa, Boulila, Ghouti, Alzahem y Latif, 2023; Qin, Zhang, Zhang, Chen, Yasunaga y Yang, 2023), por mencionar algunos ejemplos encontrados en días recientes.

En cuanto a la anotación de textos, muchos autores se encuentran a favor de su uso como herramienta automatizada (Gilardi, Alizadeh y Kubli, 2023; Huang, Kwak y An, 2023; Kuzman, Ljubešić y Mozetič, 2023), pero también los hay quienes ponen en duda su desempeño (Reiss, 2023). En el apartado de los antecedentes se profundizará este tema.

### **1.1.1 Epidemia del COVID-19**

El COVID-19 es una enfermedad causada por el virus SARS-CoV-2 (Organización Mundial de la Salud, 2021) que durante el 2020-2022 provocó confinamiento, restricciones y cuarentenas a nivel mundial desde su aparición en el 2019 (Organización Mundial de la Salud, 2022) y que tuvo sus inicios en Costa Rica en marzo del 2020 (Presidencia de la República de

---

<sup>10</sup> Disponible en <https://www.jasper.ai/chat>. [Última consulta el 28 de marzo de 2023].

<sup>11</sup> Disponible en <https://you.com/search?q=who+are+you&tbm=youchat&cfr=chat>. [Última consulta el 28 de marzo de 2023].

<sup>12</sup> Disponible en <https://writesonic.com/chat>. [Última consulta el 28 de marzo de 2023].

<sup>13</sup> Disponible en <https://www.xiaoice.com/>. [Última consulta el 20 de mayo de 2023].

<sup>14</sup> Disponible en [https://rytr.me/?via=easytouse&gclid=CjwKCAjwoIqhBhAGEiwArXT7K6j4FCM65zxoGdprGSwO\\_AOTliIAIsnOv9e8DwHoZPzHQ5VWe92-hoCkX4QAvD\\_BwE](https://rytr.me/?via=easytouse&gclid=CjwKCAjwoIqhBhAGEiwArXT7K6j4FCM65zxoGdprGSwO_AOTliIAIsnOv9e8DwHoZPzHQ5VWe92-hoCkX4QAvD_BwE). [Última consulta el 28 de marzo de 2023].

<sup>15</sup> Disponible en <https://inkforall.com/ai-safety/chatbot-fails/>. [Última consulta el 28 de marzo de 2023].

<sup>16</sup> Disponible en <https://www.perplexity.ai/>. [Última consulta el 28 de marzo de 2023].

<sup>17</sup> Disponible en <https://bard.google.com/>, aunque todavía no está abierto a la mayoría del público. [Última consulta el 28 de marzo de 2023].

Costa Rica, 2022). Aunque ya no se considera una emergencia de orden mundial (World Health Organization, 2023), los aislamientos generalizados durante el brote epidémico del COVID-19 promovieron el comercio electrónico (*e-commerce*) en Costa Rica (PROCOMER, 2021) y en el resto del mundo, porque la necesidad de virtualizar servicios presenciales de forma regular fue imperante.

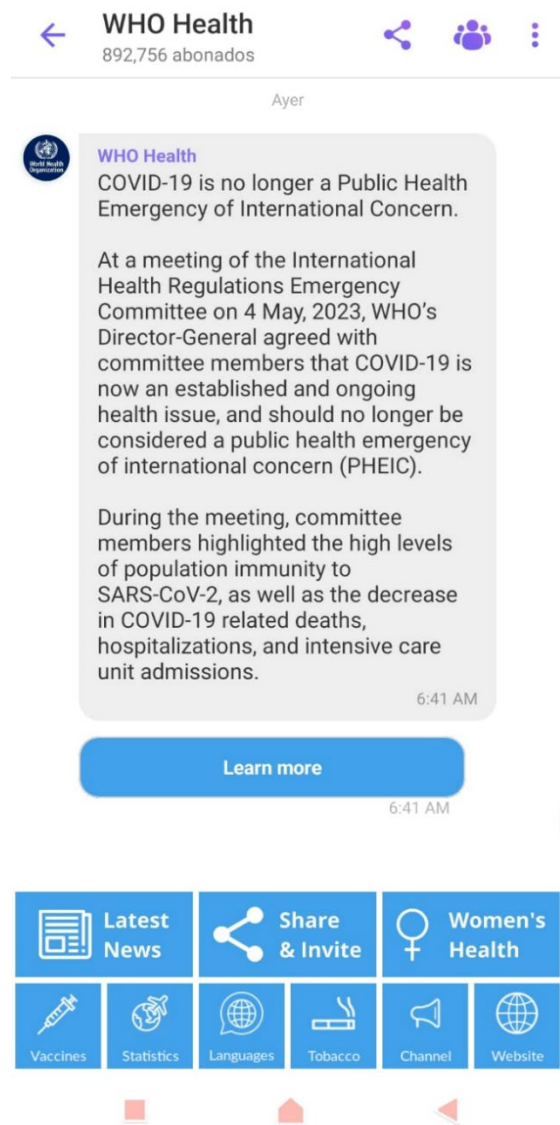
El comercio electrónico ofreció soluciones oportunas a entidades que se habían visto afectadas por la crisis sanitaria y la económica mundial y, ejemplo de esta situación son las compañías que optaron por incluir asistentes virtuales en sus páginas web como un aliado para atender al público sin tener que exponer a sus empleados o a los consumidores al contagio del virus (Comisión Económica para América y el Caribe (CEPAL), Naciones Unidas, 2020). De manera que la implementación de *chatbots* en tiempos de pandemia facilitó la comunicación entre el vendedor/proveedor y el público consumidor, en tanto que estos programas están disponibles todos los días del año y, a través de ellos, se ha contribuido al crecimiento económico local y mundial (Gupta, Hathwar y Vijayakumar, 2020).

Por otro lado, los *chatbots* se convirtieron en una herramienta de acceso directo a la información necesaria sobre la pandemia, con respuestas sobre síntomas y signos de contagio, o con recomendaciones para tratar los efectos emocionales que han causado el confinamiento y otras restricciones sociales y laborales (Miner, Laranjo y Kocaballi, 2020; Cascón-Katchadourian, 2020; Herriman, Meer, Rosin, Lee, Washington y Volpp, 2020; Neumeister, 2020; Márquez Díaz, 2020; Amer, Hazem, Farouk, Louca, Mohamed y Ashraf, 2021; Bonales Daimiel y Martínez Estrella, 2021; Fernández Moreno, 2021).

La Organización Mundial de la Salud implementó varios *chatbots* en plataformas como *Viber* (Viber Media S.à r.l., 2010-2021) y *Whatsapp* (Whatsapp Inc., 2010-2021), desde donde se ha transmitido información confiable basada en evidencia científica sobre el COVID-19 desde abril de 2020 (Jasarevic, Adhanom Ghebreyesus, Ryan, Van Kerkhove y Solomon, 2020). El *chatbot* de la Figura 1 ha sido diseñado para seguir reglas, porque responde únicamente a los comandos preestablecidos por los desarrolladores del programa. En la Figura 1, el usuario debe elegir entre los botones de abajo para seleccionar la información que desee recibir, pero no puede mantener otro tipo de conversaciones con la aplicación.

## Figura 1

### Ejemplo de Viber de WHO



La alta trascendencia de la pandemia del COVID-19 elevó el número de publicaciones sobre el COVID-19 y este exceso de información (Cascón-Katchadourian, 2020; Organización Panamericana de la Salud, 2020) ha significado que muchos estudiosos y académicos emplearan la Lingüística de Corpus y la Inteligencia Artificial, en especial el procesamiento del lenguaje natural (NLP), para compilar y analizar corpus especializados en su contexto, como en lo referente al análisis de sentimientos de usuarios de redes sociales (Kleinberg, van der Vegt y Mozes, 2020; Oyebode et al., 2020, 23 de julio; Oyebode et al., 2020, 23 de agosto; Ziems et al., 2020; Chen,

Leaman, Allot, Luo, Wei, Yan y Lu, 2021; Müller, Salathé y Kummervold, 2023) o al análisis del discurso (Bravo-Segal y Villar, 2020).

De igual manera, el procesamiento del lenguaje natural ha sido aprovechado para extraer datos y con ellos corroborar información publicada, como en el caso de los *fake news* (Organización Panamericana de la Salud, 2020); para filtrar récords médicos en el área de salud; para compilar listados de síntomas pertinentes al virus y para extraer datos de textos científicos sobre el COVID-19 (Colic, Furrer y Rinaldi, 2020; Esteva, Kale, Paulus, Hashimoto, Yin, Radev y Socher, 2020; Guo, Mirzaalian, Sabir, Jaiswal y Abd-Almageed, 2020; Lybarger, Ostendorf, Thompson y Yetisgen, 2020; Chen, Leaman, Allot, Luo, Wei, Yan y Lu, 2021), entre otros estudios.

## 1.2 Justificación

Desde el punto de vista lingüístico, el funcionamiento óptimo de un sistema de pregunta-respuesta tiene como base principal el entrenamiento por medio de corpus lingüísticos anotados con el análisis de los roles temáticos de los argumentos verbales. Los roles temáticos responden el con qué, cuándo, dónde, qué y quién de los eventos verbales, interrogantes básicas que los *chatbots* deben resolver para generar respuestas fiables a los humanos. Qué tipo de *chatbot* y qué tan inteligente se desarrolle también definirá, por supuesto, el desempeño final del sistema (Shum, He y Li, 2018; Veglis y Maniou, 2019; Gupta, Hathwar y Vijayakumar, 2020).

Otro punto importante que justifica este proyecto es que el etiquetado de roles semánticos en español continúa con numerosos problemas sin resolver en la actualidad (Ballesteros, Bohnet, Mille y Wanner, 2016; Gerdes y Kahane, 2016; Colic y Rinaldi, 2019; Nivre, de Marneffe, Ginter, Hajič, Manning, Pyysalo, S., . . . Zeman, 2020; Pyysalo, Kanerva, Missilä, Laippala y Ginter, 2015; Üstün, Bisazza, Bouma y van Noord, 2020) y estos no pueden ser relegados a sistemas de anotación automática sin la supervisión de humanos. Es por lo que, la intervención de lingüistas, con su conocimiento sobre la estructura interna de la lengua, permitirá mejorar y robustecer los modelos para el aprendizaje automático con aportes teóricos lingüísticos pertinentes.

En adición a lo anterior, es fundamental que las etiquetas de los roles temáticos que enriquecen dichos corpus estén respaldadas con análisis lingüísticos basados en teorías gramaticales que expliquen dichas relaciones sintáctico-semánticas. La investigación que se

desarrolla con esta tesis tiene relevancia académica porque aporta análisis de roles temáticos y de los argumentos verbales empleando como base teórica la Gramática Léxico-Funcional.

Desde los inicios de la AI se ha intentado agilizar la evolución de la ingeniería lingüística a través de la automatización de las tareas pertenecientes al procesamiento del lenguaje natural, prescindiendo, muchas veces, del aporte de la Lingüística (Periñán Pascual, 2012). *ChatGPT* no es la excepción a esta situación, puesto que hay quienes afirman que, en la actualidad y en un futuro cercano, los modelos de lenguaje de gran tamaño o LLM (*Large Language Model*) son o serán capaces de reemplazar a los anotadores humanos en diversas tareas, porque son menos costosos, y sus resultados son más “confiables” que el de los anotadores humanos (Gilardi, Alizadeh y Kubli, 2023; Huang, Kwak y An, 2023; Kuzman, Ljubešić y Mozetič, 2023).

En la actualidad, los usuarios de *chatbots* exigen que los sistemas de pregunta-respuesta tengan la menor cantidad de errores posibles, para que el diálogo computadora-ser humano sea más fluido y se acerque al “*imitation game*” (Coniam, 2014). El avance hacia la optimización del desempeño de los *chatbots* requiere del aporte interdisciplinar entre lingüistas e informáticos: en donde la Lingüística aporta el conocimiento relativo a la comprensión del lenguaje natural y la Informática hace este entendimiento accesible a las computadoras. Un *chatbot* debe poder responder qué, quién, cómo, cuándo, dónde y con qué y la respuesta se encuentra en las relaciones sintáctico-semánticas presentes en los argumentos y adjuntos de los eventos verbales.

La meta es que el corpus anotado final se convierta en *Gold Standard* (GSC, por sus siglas del inglés *Gold Standard Corpora*) y que llegue a ser explotado por otras áreas académicas en el procesamiento del lenguaje natural, particularmente en la creación de sistemas de pregunta-respuesta, para validar la calidad de dichos sistemas. Los GSC son colecciones estándar de textos anotados, los cuales han sido revisados y anotados por expertos en el tema y el acuerdo entre los anotadores ha sido computado para asegurar la calidad final del corpus (Wissler, Almashraee, Monett y Paschke, 2014).

Las razones anteriores justifican la importancia de desarrollar un modelo de anotación en español para los roles temáticos en donde, tanto anotadores humanos como un modelo entrenado con AI, en particular *ChatGPT*, asignen estas etiquetas para contrastar el desempeño de ambas partes. Este modelo de anotación y el corpus anotado final serán de gran apoyo lingüístico para fortalecer los programas de aprendizaje automático creados por otras academias. El trabajo conjunto con métodos analizados desde la Lingüística permitirá disponer de recursos

lingüísticamente más precisos y de *chatbots* que produzcan respuestas en español más acertadas y oportunas en el futuro.

### 1.3 Planteamiento del problema

Los corpus se deben anotar con descripciones lingüísticas para que puedan ser empleados en las aplicaciones de NLP, en *Machine Learning* y, principalmente, para el entrenamiento de sistemas de pregunta-respuesta. El mayor problema que se encuentra al momento de llevar a cabo esta investigación es que no hay corpus sobre el COVID-19 que hayan sido analizados sintácticamente, ni anotados con los roles temáticos de los argumentos verbales. Mucho menos los hay que tengan el estudio detallado para responder las preguntas comunes en español.

A pesar de que sí existen numerosos corpus compilados sobre la pandemia, como el de Brosa Rodríguez (2022) desde la perspectiva sociolingüística, no los hay con los análisis lingüísticos y sus anotaciones pertinentes para que equipos de otras academias, como los del área de Informática, puedan explotarlos en la creación de *chatbots*. Esta exigüidad dificulta la creación, entrenamiento y aplicación de *chatbots* sobre esta temática en particular, y es también un impedimento para fortalecer los estudios de NLP en español.

La presente tesis se enfoca en crear un modelo de anotación de roles temáticos en español que pueda ser explotado para el entrenamiento de *Machine Learning*. La anotación de roles temáticos en español es un tema que ha permanecido al margen, puesto que muchas corrientes teóricas han dado prioridad a otros aspectos de las lenguas naturales. Estos vacíos en NLP obligan a la academia lingüística a llevar a cabo una contribución válida en el estudio de corpus lingüísticos y al etiquetado de roles temáticos en la lengua española. Máxime si este aporte se da con análisis desarrollados y explicados desde una teoría como la Gramática Léxico-Funcional, creada y pensada para la anotación de corpus lingüísticos.

Otro problema lingüístico es la carencia de estudios sobre *chatbots* de tercera generación en español, como *ChatGPT*, en cuanto a su desempeño como anotadores automatizados de etiquetas lingüísticas. Los *chatbots* de tercera generación cuentan con modelos lingüísticos completos que les permite generar respuestas personalizadas y aprender de las conversaciones sin entrenamiento previo (*Generalized Zero-Shot Learning Methods, GZSL* por sus siglas en inglés) (Pourpanah, Abdar, Luo, Zhou, Wang, Lim, . . . Wu, 2023).

Este tipo de *chatbots* ofrece resultados excepcionales en diversas tareas, como las relacionadas con respuesta a preguntas, traducción, generación de texto o en resúmenes automáticos (Zhang, Ding y Jing, 2022; Qin, Zhang, Zhang, Chen, Yasunaga y Yang, 2023; Zhong, Ding, Liu, Du y Tao, 2023; Peng, Ding, Zhong, Shen, Liu, Zhang, ... y Tao, 2023; Lee, 2023, 20 de abril), porque poseen mayor capacidad para recordar el contexto de las conversaciones a largo plazo y aprender de ellas (Pérez García, 2021). No obstante, sus respuestas no deben tomarse como válidas a ciegas, sino que humanos expertos en el tema deben evaluarlas siempre.

En adición a los problemas aludidos en párrafos anteriores, los estudios que se presentarán en los antecedentes comprueban que es imprescindible una revisión manual llevada a cabo por lingüistas, porque la automatización de los etiquetados morfológico, morfosintáctico, sintáctico y semántico no carece de errores e incoherencias. Para conseguir un entramiento automático eficaz, se requiere de un corpus con etiquetas precisas, basadas en gramáticas pertinentes para la lengua española sin traducciones o montajes *ad hoc* de ontologías de otras lenguas, como del inglés, por ejemplo.

De los párrafos anteriores se desprende que es esencial que los análisis de los roles temáticos y de los argumentos verbales, fundamentales para el aprendizaje automático y el entrenamiento de *chatbots*, se sirvan sí de la automatización para ciertas tareas de NLP, como *parsing* y etiquetado de roles temáticos, pero que los resultados de esta automatización sean siempre comprobados con teorías lingüísticas pertinentes para que los corpus en español que se creen estén enriquecidos con información válida.

#### 1.4 Objetivos de la tesis

En esta tesis se busca crear un modelo de anotación de roles temáticos en español, específicamente para sistemas de pregunta-respuesta, empleando la Lingüística de Corpus como metodología y la Gramática Léxico-Funcional (LFG) como base teórica. En concreto, el objetivo principal para esta tesis es:

- ❖ Desarrollar una propuesta de modelo de anotación de roles temáticos en español específico para sistemas de pregunta-respuesta.

Para llevar a cabo este objetivo, se proponen los siguientes objetivos específicos:

- A. Crear un subcorpus con los 200 verbos más frecuentes del *Corpus COVID-19*.
- B. Clasificar los argumentos verbales y adjuntos de las frases del subcorpus en cuanto a sus Funciones Temáticas (agente, meta, experimentador, instrumento, tema, fuente, lugar y tiempo) y a sus Funciones Gramaticales, para constatar las relaciones sintáctico-

- semánticas que responden las interrogantes qué, cómo, cuándo, dónde, quién, con qué y a quién.
- C. Desarrollar una guía de anotación para el modelo de etiquetado de los roles temáticos presentes en el corpus para que los anotadores sigan las instrucciones efectivamente.
  - D. Evaluar el acuerdo entre los anotadores humanos por medio del coeficiente Kappa de Fleiss.
  - E. Evaluar el desempeño de *ChatGPT* como anotador de los roles temáticos definidos para esta investigación en contraste con el de los anotadores humanos por medio del coeficiente Kappa de Fleiss.
  - F. Comparar el desempeño de los anotadores humanos frente al de *ChatGPT*, tomando como punto de partida el corpus anotado manualmente por la tesiaría, por medio del coeficiente Kappa de Fleiss.

### 1.5 Alcances y limitaciones de la tesis

Dentro de los alcances de la tesis se buscó llevar a cabo un análisis exploratorio y descriptivo de los niveles sintáctico y semántico para los 200 verbos más frecuentes del *Corpus COVID-19*, con el fin de anotar los roles temáticos que responden las interrogantes de los sistemas de pregunta-respuesta: agente, meta, experimentador, instrumento, tema, fuente, lugar y tiempo, por parte tanto de anotadores humanos como de *ChatGPT*.

La tesiaría consideró el carácter exploratorio como pertinente para la anotación de los roles temáticos, por tratarse de una investigación empírica con pocos antecedentes para la lengua española, como se verá más adelante, y por emplear la LC como método de análisis. Los resultados describen el fenómeno lingüístico que ocurre dentro de este corpus, en particular en relación con los roles temáticos de los argumentos y adjuntos del verbo; además, se deja una guía de anotación que puede ser empleada en futuros trabajos de Lingüística Computacional para anotadores humanos y para LLMs.

Una limitante importante para realizar esta investigación ha sido la carencia de trabajos anteriores sobre el etiquetado de los roles temáticos para *chatbots* en español, de los que se pudieran tomar parámetros y ejemplos de metodología a seguir, principalmente estudios sobre *ChatGPT* publicados por lingüistas. En consecuencia, esta tesis se enfocó en el área meramente lingüística de la anotación y no se entró en detalles en cuanto a la creación en sí de los *chatbots* ni

a su entrenamiento automático, como tampoco se evaluó el modelo de anotación con algoritmos computacionales para este propósito, porque, principalmente, todo lo anterior se sale del campo académico de esta investigación.

Otra limitante fue encontrar un balance en la cantidad de ejemplos para cada categoría de los roles temáticos establecidos para responder las preguntas de los *chatbots*. Se subraya este objetivo, porque, como se trabaja con un corpus que emplea el uso natural de la lengua en español y por ser esta una investigación empírica, se corría el riesgo de no obtener la misma cantidad de casos para cada categoría de los roles definidos.

En cuanto a los análisis sintáctico y semántico, se limitó a categorizar los argumentos según su función gramatical y su función semántica. No se entró en detalles con respecto a los rasgos morfológicos como género, número, tiempo verbal, aspecto o modo. Sin embargo, todas estas características fueron tomadas en cuenta para estudiar cada frase y para categorizar sus componentes correctamente. La razón por la cual se llegó a esta decisión es porque, de haberse realizado un estudio morfológico más profundo, no se habría beneficiado la anotación de los roles temáticos en sí y se habría extendido innecesariamente el proceso investigativo y el de anotación.

Finalmente, no siempre es factible encontrar suficientes anotadores entrenados en el fenómeno lingüístico y que, además, tengan la disponibilidad para participar en una investigación de considerable tamaño. De manera que, se trabajó con la mayor cantidad de anotadores disponibles en el momento para probar el modelo de anotación. En cuanto a las discordias encontradas entre los anotadores, la tesiaría asignó un tercer anotador que no había tenido parte en el modelo anteriormente para que resolviera el desempate.

En este apartado se detallaron la contextualización del tema, la justificación de la investigación, el planteamiento del problema, los objetivos, y el alcance y las limitaciones con que se llevó a cabo este trabajo. En 1.6 se introducirá el contexto y los antecedentes al corpus lingüístico, particularmente para los corpus anotados con los roles temáticos en español. En 1.7 se expone el marco teórico, mientras que en 1.8 se define la metodología para alcanzar los objetivos precisados en este capítulo.

## 1.6 Contexto y antecedentes a corpus lingüísticos

En este apartado se presenta el contexto del que parte la anotación de corpus lingüísticos y los antecedentes que motivaron esta investigación académica. La distribución de los temas dentro de este apartado seguirá el siguiente orden. En 1.6.1 se describe un breve contexto del procesamiento del lenguaje natural como método de extracción de la información; en 1.6.2, sobre el enfoque simbólico; en 1.6.3 sobre corpus y LC; en 1.6.4 se detallarán los antecedentes al corpus en español y en 1.6.5, los antecedentes del corpus en español con anotaciones semánticas. En 1.6.6 se dedicará un apartado sobre LLM y en 1.6.7 se expone el estado de la cuestión.

### *1.6.1 Procesamiento del lenguaje natural como método de extracción de la información*

El procesamiento del lenguaje natural (NLP, por sus siglas en inglés) deriva del estudio del lenguaje natural y de la Lingüística Computacional (en adelante LComp, para diferenciarla de LC), la cual es parte de la AI y de la estadística (Periñán Pascual, 2012). Es común en la literatura encontrar el uso intercambiable entre los términos NLP y LComp (Llisterri, 2003; Navarro-Colorado, 2007; Periñán Pascual, 2012; Pustejovsky y Stubbs, 2013), porque en sus inicios, la LComp era la ‘lingüística para las computadoras’ y tenía como tarea principal proveer de la capacidad de entender el lenguaje natural a la computadora (Gelbukh, 2010).

Sin embargo, es preciso señalar que la mayor diferencia entre LComp y NLP se basa en que la LComp sostiene un punto de vista enfocado más en lo teórico, según Llisterri (2003), puesto que esta busca la modelización del conocimiento lingüístico con el objetivo de construir sistemas computacionales que sean capaces de analizar y generar textos en lenguaje natural; mientras que el NLP, por medio de aplicaciones concretas, busca dar soluciones a los problemas que la LComp plantea con sus teorías (Periñán Pascual, 2012). En adelante, para la presente investigación, se hará referencia a NLP como término general, tanto para la teoría como para la aplicación de sus técnicas.

Como campo interdisciplinario que es, el NLP incorpora los conocimientos y esfuerzos conjuntos de la informática, de la lingüística, de la ciencia cognitiva y de diversas ingenierías, con la finalidad de estudiar la interacción hombre-máquina (Garbade, 2018), así como el diseño y la creación de aplicaciones para mejorar dicha interacción a través del lenguaje humano (Pustejovsky y Stubbs, 2013). En palabras de Michael J. Garbade: «El objetivo final del PNL es leer, descifrar,

comprender y dar sentido a los lenguajes humanos» (2018, párr. 6)<sup>18</sup>, para lo cual se busca convertir el lenguaje natural en un lenguaje formal (Salguero Lamillar, 1995), esto es, en un lenguaje que las computadoras puedan comprender, procesar y generar, como ocurre con los lenguajes de programación computacional (Instituto de Ingeniería del Conocimiento, s.f.).

Durante los años 60 y 70, los investigadores se dividieron en dos ramas de estudio de NLP: la del enfoque simbólico y la del enfoque estocástico o probabilístico, con el propósito de dotar las computadoras con herramientas y conocimiento para que fueran capaces de comprender el lenguaje natural. Los investigadores del enfoque simbólico (o basado en reglas) consideraban que esta rama fue el inicio de la investigación de la AI, puesto que buscaba dar solución a los problemas teóricos más significativos del NLP, por medio de la creación de sintaxis y de lenguajes formales (Gallagher, Rafferty y Wu, 2004; Perrián Pascual, 2012).

Por su lado, el enfoque probabilístico se preocupaba más por métodos estadísticos y en el reconocimiento de patrones entre textos (Gallagher, Rafferty y Wu, 2004). El enfoque probabilístico tuvo su auge «en las tecnologías del habla, el etiquetado gramatical, el análisis sintáctico y la semántica» (Perrián Pascual, 2012, pág. 24). Perrián Pascual (2012) opina que los enfoques estadísticos han servido como complemento a los enfoques simbólicos para comprobar la validez de los argumentos que académicos, como Noam Chomsky, entre otros, habían propuesto, y se ha descubierto que, a la hora de aplicar dichos argumentos, estos no eran empíricamente precisos y que, por lo tanto, el enfoque simbólico empleado requería una revisión más profunda para mejorar su exactitud (Gallagher, Rafferty y Wu, 2004).

Un ejemplo en práctica de estos dos enfoques se aprecian en SHRDLU y PARRY, dos sistemas basados en diálogo: SHRDLU empleaba un análisis lingüístico que tenía como base un modelo teórico gramatical, así como un razonamiento basado en la lógica que incorporaba conocimiento de mundo; mientras que PARRY, al igual que ELIZA, aplicaba un reconocimiento simple de patrones en la superficie de la entrada textual y, por medio de un módulo de interpretación-acción, PARRY podía recrear respuestas similares a las de un paciente con paranoia (Perrián Pascual, 2012).

En sus inicios, la AI tenía como objetivo automatizar tareas que habían sido ejecutadas manualmente por el ser humano, como la traducción textual. No obstante, los intentos fallidos por

---

<sup>18</sup> Traducción personal de “The ultimate objective of NLP is to read, decipher, understand, and make sense of the human languages in a manner that is valuable”.

traducir automáticamente mostraron un vacío que debían resolver: necesitaban enseñar la lengua natural a las computadoras para que estas aprendieran a procesar el lenguaje de los humanos. Paralelamente al crecimiento en la investigación en IA, los investigadores de NLP buscaban desarrollar gramáticas con el enfoque simbólico, con lo cual, se pensaba, se solucionarían los problemas semánticos y de ambigüedad que obtenían con los primeros programas de procesamiento de lenguaje natural (Periñán Pascual, 2012).

### ***1.6.2 El enfoque simbólico***

Con Noam Chomsky nació lo que hoy se conoce como la Gramática Generativa Transformacional (GGT) o ‘*mainstream generative grammar*’ (o modelo generativo). Chomsky desarrolló sus estudios sobre la sintaxis del lenguaje empleando herramientas de origen matemático, pero no con la intención de que fueran aplicados en procesos computacionales (Chomsky, 1987; Luque, 2017). Su gramática se basaba en los procesos mentales<sup>19</sup> de los hablantes, en donde la semántica y la fonología se podían relegar a componentes interpretativos, y como una capa más de la sintaxis, la cual era el eje de su gramática (Ozkural, 2017).

Los investigadores de NLP se sintieron atraídos por la GGT, porque era un modelo formal del lenguaje natural y este facilitaba su implementación en las computadoras, al menos desde el punto de vista teórico (Periñán Pascual, 2012). Con su aplicación todavía quedaban sin resolver los problemas semánticos y los de ambigüedad, por lo que el interés en la GGT se fue perdiendo durante la década de 1970 conforme se profundizaban los estudios sobre semántica (Periñán Pascual, 2012).

En un intento por integrar la semántica en los modelos generativistas, se crearon las teorías lexicistas en los años 80. En ellas se da mayor importancia al léxico y menor al papel de la sintaxis. Ejemplo de estas teorías son la Gramática Léxico-Funcional, en adelante LFG por sus siglas en inglés (Kaplan y Bresnan, 1982; Bresnan, 2001; Bresnan, Asudeh, Toivonen y Wechsler, 2016), la Gramática de Estructura Sintagmática Generalizada (Gazdar, Klein, Pullum y Sag, 1985), la Gramática de Estructura Sintagmática Nuclear (Pollard y Sag, 1987) y la Teoría Sentido-Texto (Mel'cuk, 1988, citado en García Marchena, 2009).

---

<sup>19</sup> Sobre los procesos mentales, la Gramática Universal y discusiones al respecto, ver Fodor, 1986; Chomsky, 1987; Petro, 1996; Alonso-Cortés, 2004; Siegel, 2007; Pinker, 2009; Berwick, 2009; Rice, 2011; Dehaene, 2015.

El paradigma simbólico vino a solucionar parte de los problemas que el NLP intentaba resolver con una forma de lenguaje formal que las computadoras podían comenzar a interpretar. Sin embargo, el uso del modelo simbólico presentaba múltiples limitaciones, al igual que el estocástico, como la falta de flexibilidad necesaria para tratar adecuadamente *inputs* o entradas con oraciones agramaticales, con elipsis o con realizaciones lingüísticas nuevas (Periñán Pascual, 2012). En otras palabras, las reglas del lenguaje formal basadas en el enfoque simbólico no daban cuenta de las excepciones y de los cambios que las lenguas naturales tienen y que son común en los hablantes. Las gramáticas formales del modelo simbólico se enfocaban en estándares de la lengua y no en el uso real de ella, de aquí que se tuviera que cambiar el rumbo de la investigación lingüística para dar cuenta de la realidad del hablante.

### ***1.6.3 Corpus y Lingüística de Corpus***

En la década de los 80 comenzaron las investigaciones enfocadas en *Speech Recognition* y, con estas, se compilaban datos para crear modelos de lenguaje utilizando *n-grams* y los Modelos Ocultos de Markov (MOM o HMMS en inglés), con los que se podían reconocer un limitado conjunto de palabras en un dominio reducido (Pustejovsky y Stubbs, 2013). Con la llegada del Internet en los 90 se obtuvo acceso a recursos lingüísticos textuales y con los grandes avances computacionales, como mayor velocidad para procesar datos y el aumento en el espacio de memoria, se comenzaron a compilar y almacenar corpus extensos con usos actuales en diversos idiomas, para trabajar el NLP hacia la ingeniería lingüística (Periñán Pascual, 2012; Pustejovsky y Stubbs, 2013).

Aunado a esto y a las limitaciones que presentaba el modelo simbólico, resurgieron las tendencias de orden empirista enfocadas en la aplicación de métodos estocásticos para el procesamiento del lenguaje natural, principalmente, como complemento de los modelos simbólicos en la década de los 90. Los modelos empiristas dan auge al modelado estadístico del lenguaje para la traducción y para trabajar en *Machine Translation* (Periñán Pascual, 2012; Pustejovsky y Stubbs, 2013).

El componente estructuralista de J.R. Firth tuvo mucho peso en el paso del modelo simbólico al estudio estocástico por medio de los corpus lingüísticos. En palabras de Rojo (2021), J.R. Firth afirmaba que para conocer realmente el significado de una palabra se debía analizar *todo* el contexto dentro del cual se empleaba dicha palabra, de aquí que considerara los corpus como un recurso que facilitaba la investigación de los aspectos relevantes para comprender los distintos

fenómenos lingüísticos presentes en las lenguas. Guillermo Rojo (2021) explica que el componente estructuralista de Firth supuso grandes cambios con respecto a los enfoques tradicionales, no solo de la gramática sino también del léxico. Entre estos cambios, los corpus comprobaron ser la mejor manera de exponer la configuración estadística de los textos, y su estudio facilita observar y estudiar las frecuencias y dispersiones léxicas presentes en los textos.

Los corpus también permiten abordar adecuadamente todo lo relacionado con fraseología y con el análisis de las colocaciones o coapariciones. Guillermo Rojo (2021) resalta que uno de los mayores aportes de la creación y el estudio de corpus es la codificación de los textos que los componen. A partir de esta codificación, llamadas también ‘anotaciones’, es posible crear y recuperar corpus virtuales, desde donde se pueden seleccionar fenómenos lingüísticos para estudios comparativos de una variedad, ya sean estos para verificar aspectos diacrónicos, diatópicos o diastráticos, por ejemplo.

En las décadas anteriores no se habían aplicado métodos estadísticos y cuantitativos en gran escala porque los investigadores debían trabajar con materiales lingüísticos de tamaño reducido y esto imposibilitaba llevar a cabo generalizaciones estadísticas que aportaran análisis confiables e interesantes sobre la lengua que estudiaban (Pustejovsky y Stubbs, 2013). Además, los partidarios de la investigación lingüística de vanguardia durante la década de los 60, es decir, la gramática generativo-transformacional, consideraban el corpus como un recurso inútil, porque lo relacionaban con el concepto empleado, hasta ese momento, por los seguidores de Leonard Bloomfield y el distribucionalismo (Rojo, 2021).

Para la década de 1950, los lingüistas descriptivos habían compilado corpus con colecciones de expresiones orales y escritas, como el KWIC (*Key Word in Context*) para indexar documentos y crear concordancias (Pustejovsky y Stubbs, 2013). Entre 1963 y 1964 se creó el primer corpus general, moderno y legible para computadoras, el *Brown Corpus of Standard American English (Brown Corpus)*, de la Universidad de Brown (University of Essex, 1998; Pustejovsky y Stubbs, 2013; Rojo, 2016).

Durante esta misma década se desarrollaron técnicas para trabajar con similitud estadística en el contenido de documentos, dentro de la ciencia que se conoce como Búsqueda y Recuperación de Información, en inglés *Information Search and Retrieval (ISR)* (Pustejovsky y Stubbs, 2013). Hacia los años 70, la Lingüística de Corpus, junto con los modelos estocásticos, crearon los

sistemas Reconocimiento de Voz (*Speech Recognition*) y con el *Survey of English Usage* se desarrolla el *London-Lund Corpus (LLC)* (Pustejovsky y Stubbs, 2013).

En los 80, el *Survey of English Usage* da lugar al *Grammar of English* y al *Lancaster-Oslo-Bergen (LOB)*, diseñado para igualar al *Brown Corpus*. El *Bank of English*, publicó el diccionario COBUILD (*Collins Birmingham University International Language Database*) y el *Child Language Exchange System (CHILDES)*, un repositorio de datos en la adquisición del primer idioma (Pustejovsky y Stubbs, 2013).

En los 90 se destacan el *British National Corpus (BNC)*, el corpus más grande en inglés al momento con 100 millones de palabras y el *Penn TreeBank*, el primer *treebank* con oraciones que ocurren naturalmente en inglés, etiquetadas y analizadas sintácticamente (Pustejovsky y Stubbs, 2013). A diferencia de los corpus, un *treebank* es «un segmento de un corpus que ha sido enriquecido con información lingüística suplementaria» (Jara Murillo, 2013), con «conjuntos de secuencias analizadas sintácticamente en los que es posible recuperar las que reúnen unas determinadas condiciones estructurales» (Rojo, 2016, p. 292).

Ya en el siglo XXI, con el crecimiento del *World Wide Web* se ponen a disposición más datos con los que se pueden ejecutar modelos estadísticos para *Machine Translation* (Pustejovsky y Stubbs, 2013). En la primera década de este siglo se publican *American National Corpus (ANC)*, *Corpus of Contemporary American English (COCA)* y *Google N-gram Corpus*; y, hacia la segunda década, se ponen a disposición nuevas formas de corpus textuales recolectadas en las redes sociales, con las cuales se crean modelos para analizar toda una nueva gama de fenómenos lingüísticos (Pustejovsky y Stubbs, 2013).

Siguiendo la explicación de Guillermo Rojo (2021), los corpus son muestras representativas de un universo lingüístico que permiten analizar las producciones lingüísticas efectivamente realizadas por los hablantes, y sus resultados producen materiales apropiados para la investigación científica. Guillermo Rojo propone que «los corpus constituyen el modo más adecuado de llevar a cabo estudios de carácter empírico (con datos objetivos, replicables y pertinentes con relación al fenómeno estudiado)» (2021, pág. 1), porque son un «conjunto estructurado, elaborado de forma sistemática, organizado de una determinada forma y con una cierta relación cuantitativa precisa entre los subcorpus que lo componen» (2021, pág. 3).

El corpus, como conjunto estructurado, debe contener entonces textos que sean pertinentes para la temática que se tenga interés en abordar. Para esta tesis, la pandemia del COVID-19 es el

foco principal sobre el que se estudiarán las relaciones sintáctico-semánticas de los argumentos verbales y los adjuntos. Además, debe ser elaborado de forma sistemática, de manera que los textos que lo compongan sean todos realizados en la misma lengua o variedad lingüística, es decir, en el mismo tipo de texto (oral o escrito), en el mismo formato, organizado de una determinada forma con las anotaciones pertinentes para su análisis, y con una cierta relación cuantitativa precisa que dé cuenta de todos los casos del fenómeno que se indagará.

Los corpus muestran las características globales de la población textual que representa y de estas se pueden realizar análisis científicos (Rojo, 2021). Existen diversos tipos de corpus y su configuración dependerá de los objetivos para los cuales se compilen. Desde la posición de Guillermo Rojo (2016; 2021), los corpus pueden ser sincrónicos o diacrónicos, generales, de referencia o especializados. Ejemplos de estos son los que se han dado en español, los cuales tuvieron sus inicios entre 1980 y 1995, con la aplicación de la Lingüística de Corpus como su método de análisis, (Parodi, 2010; Rojo, 2016; Rojo, 2021). Los corpus en español se han caracterizado por haber sido creados dentro «de un marco tecnológico más evolucionado», en comparación con el *Brown Corpus* (Rojo, 2016, p. 287), porque, desde sus inicios, se concibieron para que fueran accesibles a un mayor número de investigadores.

Además, se caracterizaron por ser de tamaño reducido y emplearse como fuentes de datos para análisis variados (*Lovaina*, *ENTREVIS90*, *ENTREVIS95*, *Spanish online (SOL)*, *Corpus lingüístico de referencia de la lengua española en Argentina*, el *Corpus lingüístico de referencia de la lengua española en Chile*, *LEXESP* y *Corpus of Contemporary Spanish*), así como también para propósitos lexicográficos (*Corpus Vox-Biblograf (CVB)*, *CUMBRE*, *Corpus del español mexicano contemporáneo (CEMC)*) y para estudios diacrónicos (*ADMYTE*) (Rojo, 2016).

Dentro de los corpus especializados, están los corpus de entrenamiento que son indispensables en el procesamiento automático de grandes cantidades de texto, a los cuales se les da el nombre de *gold standard*. Existen corpus en español sobre temas especializados, entre ellos, el *Iberia*, que fue creado en 2008-2009 y trabaja con textos científico-técnicos que datan entre 1985 y 2015 (Rojo, 2016), y todas sus palabras han sido indexadas y analizadas lingüísticamente, aunque carece de anotaciones semánticas (Grupo de investigación ELCI, s.f.). Este corpus puede accederse públicamente por medio de su página web<sup>20</sup> para realizar consultas únicamente. El

---

<sup>20</sup> Disponible para consulta en <http://iberia.cchs.csic.es/cgi-bin/iberia.cgi>. [Última consulta el 31 de marzo de 2023].

proyecto *PANACEA*<sup>21</sup> se dedica a textos relacionados con el medio ambiente y la legislación laboral (Rojo, 2016). *PANACEA* se enfoca en la creación de recursos lingüísticos para adquisición, producción, actualización y mantenimiento de datos para sistemas de traducción automática (Bel, 2010).

Otra característica de los corpus es que estos pueden ser orales o escritos, en cuanto al soporte material del texto. Es necesario agregar que, en un principio, los corpus orales tendían a ser más coloquiales, mientras que los escritos eran más literarios; no obstante, los blogs y las redes sociales han dado lugar a nuevos géneros, en donde el soporte escrito se acopla con la lengua coloquial (Rojo, 2021). Ejemplos de corpus en español, tanto diacrónicos como sincrónicos, son el Corpus Histórico del Español en México CHEM, el Corpus de Documentos Españoles Anteriores a 1800 CODEA, Corpus Hispánico y Americano en la Red: Textos Antiguos CHARTA, Corpus Diacrónico y Diatópico del Español de América CORDIAM, el Corpus Escrito del Español como L2 CEDEL2, el Corpus de aprendices de español CAES del Instituto Cervantes, entre otros.

Una particularidad importante de los corpus es que pueden ser cerrados o abiertos, en donde los primeros no admiten más datos y tienen caducidad, mientras que los corpus abiertos, reciben e incorporan textos nuevos, cambian de contenido con el tiempo y los resultados de sus estudios se modifican cada vez que se lleva a cabo un nuevo análisis. Ejemplos de corpus cerrados en español son la *Base de datos sintácticos del español actual*<sup>22</sup> (BDS), de la Universidad de Santiago de Compostela, que como corpus está actualmente cerrada y, la mayoría de sus datos provienen de obras literarias pertenecientes al siglo anterior. A partir de esta base de datos se han desarrollado otros corpus, como el CSA, del proyecto DRASAE, o la base de datos ADESSE<sup>23</sup>, en la Universidad de Vigo (Vaamonde, 2008). Otro ejemplo es *Internet corpora*<sup>24</sup>, construido por la Universidad de Leeds (Sharoff, n.d.), que permite realizar *queries* sobre concordancias y colocaciones sobre 143 millones de formas (Rojo, 2016) en Internet (Sharoff, n.d.); sin embargo, la mayoría de las páginas web que revisa están descontinuadas o fuera de servicio.

Continuando, los corpus pueden ser monolingües o multilingües, en donde los últimos son más empleados en traducciones y en estudios comparativos entre lenguas (Rojo, 2021). En español

---

<sup>21</sup> Ver <http://www.panacea-lr.eu/>. [Última consulta el 31 de marzo de 2023, pero la página web ya no estaba funcionando.

<sup>22</sup> Disponible para consulta en <http://www.bds.usc.es/>. [Última consulta el 31 de marzo de 2023].

<sup>23</sup> Disponible para consulta en <http://adesse.uvigo.es/data/>. [Última consulta el 31 de marzo de 2023].

<sup>24</sup> Disponible para consultas en <http://corpus.leeds.ac.uk/internet.html>. [Última consulta el 31 de marzo de 2023].

existen algunos que forman parte de corpus multilingües, como el *Catalan-Spanish Parallel Corpus (ELRA-W0053)*; el *European Parliament proceedings parallel corpus 1996–2011* y el *Wikicorpus*<sup>25</sup> v. 1.0 (Reese, Boleda, Cuadros, Padró y Rigau, 2010), con anotaciones hechas con *FreeLing* (Rojo, 2016).

Existen otros tipos de corpus, los llamados “oportunistas”, los cuales se obtienen a partir de textos públicos en la red y ofrecen la facilidad de su obtención y la limitación de los costes de creación. Sin embargo, los corpus oportunistas tienen una codificación deficiente y de difícil extracción en algunos casos, además de que la variedad de textos que se consiguen por este medio es limitada, porque, normalmente, provienen de blogs y de páginas web (Rojo, 2021). En español se han producido varios corpus basados en Internet. Uno de ellos es el *Internet corpora*, anteriormente mencionado, y el *Web as Corpus*<sup>26</sup> (Research and Development Unit for English Studies, 1999) que se caracteriza por ser más un archivo y no tanto un corpus en el sentido estricto de la palabra, porque se obtiene poca información lingüística y esto lo convierte en una herramienta poco selectiva para investigaciones especializadas para el procesamiento del lenguaje natural (Rojo, 2016).

Finalmente, el corpus *EsTenTen*<sup>27</sup> que, al 2018, contenía 17.5 mil millones de palabras procedentes de la web europea y americana en español, y de la totalidad de Wikipedia en español (Lexical Computing, 2020), cuenta con etiquetado gramatical (*part-of-speech tagging*, *POS tagging* o *POST*) y con lematización llevados a cabo por medio de *FreeLing* (Lexical Computing, 2020). Su acceso es por medio de suscripción pagada.

#### ***1.6.4 Antecedentes al corpus en español***

A partir de 1995, la Real Academia Española (RAE) crea varios corpus de referencia, los cuales no cuentan con anotaciones lingüísticas, porque no fueron diseñados para el procesamiento del lenguaje natural (Rojo, 2016). Ejemplos de estos son el *Corpus de referencia del español actual*<sup>28</sup> (CREA), el *Corpus diacrónico del español*<sup>29</sup> (CORDE), el *Corpus del diccionario*

---

<sup>25</sup> Disponible para descarga en <https://www.cs.upc.edu/~nlp/wikicorpus/>. [Última consulta el 31 de marzo de 2023].

<sup>26</sup> Disponible para consultas en <http://www.webcorp.org.uk/live/>. [Última consulta el 31 de marzo de 2023].

<sup>27</sup> Disponible en <https://www.sketchengine.eu/estenten-spanish-corpus/>, por medio de suscripción pagada. [Última consulta el 31 de marzo de 2023].

<sup>28</sup> Disponible para consultas en <http://corpus.rae.es/creanet.html>. [Última consulta el 31 de marzo de 2023].

<sup>29</sup> Disponible para consultas en <http://corpus.rae.es/cordenet.html>. [Última consulta el 31 de marzo de 2023].

histórico<sup>30</sup> (CDH) y el *Corpus del español del siglo XXI*<sup>31</sup> (CORPES), estos dos últimos como bases para diccionarios (Alcántara-Plá, s.f.; Rojo, 2016).

El *Corpus del español*<sup>32</sup> (CE), de Mark Davies, «está parcialmente anotado y lematizado» (Rojo, 2016, p. 289), pero tiene un carácter referencial y no está diseñado para NLP. Otros corpus de carácter general son el *Corpus del español actual*<sup>33</sup> (CEA) (Alcántara-Plá, s.f.; Rojo, 2016), el cual cuenta con una interfaz pública para consultas en línea. El *Corpus dinámico del castellano de Chile*<sup>34</sup> (Codicach), de Scott Sadowsky (Sadowsky, 2006; Rojo, 2016), ha sido «chunkeado, lematizado, y etiquetado con información sobre categorías gramaticales y relaciones sintácticas tanto con *FreeLing* como con *Machinese Syntax de Connexor*» (Sadowsky, 2006, párr. 1); no obstante, este corpus requiere permiso específico de su autor para poder consultarse.

De un tamaño más pequeño, se encuentra el *Corpus Ancora-Es*<sup>35</sup> ((Taulé, Martí y Recasens, 2008; Vaamonde, 2008), del grupo CLIC, de la Universidad de Barcelona, el cual contiene información léxica y gramatical (Rojo, 2016). *Corpus Ancora-Es* ha sido analizado en diversas áreas, en particular en la anotación de argumentos implícitos de los nombres deverbales (Peris Morant y Taulé Delor, 2011; Taulé, Peris y Rodríguez, 2016). Por su lado, el corpus *GRIAL*<sup>36</sup> estudia las diferencias entre registros del español (Parodi, 2006) y ofrece al público una interfaz computacional para anotaciones morfosintácticas, e interrogaciones de corpus en español (ver más adelante el estudio realizado por Vázquez, Alonso, Capilla, Castellón y Fernández (2006).

En cuanto a los *treebanks*, los más destacados para el español son el *RST Spanish treebank*<sup>37</sup> de la Universidad Autónoma de México; el *Spanish tree-bank*<sup>38</sup> de la Universidad

---

<sup>30</sup>Disponible para consultas en <https://webfrrl.rae.es/CNDHE/view/inicioExterno.view;jsessionid=C36320ADED63B644B8AC44FD31EB194E>. [Última consulta el 31 de marzo de 2023].

<sup>31</sup> Disponible para consultas en <https://webfrrl.rae.es/CORPES/view/inicioExterno.view>. [Última consulta el 31 de marzo de 2023].

<sup>32</sup> Disponible para consultas en <https://www.corpusdelespanol.org/>. [Última consulta el 31 de marzo de 2023].

<sup>33</sup> Disponible para consultas en <http://cea.spanishfn.org/cea/>. [Última consulta el 31 de marzo de 2023].

<sup>34</sup> No está disponible en línea, pero se puede solicitar el acceso por medio de su autor (Sadowsky, 2006).

<sup>35</sup> Puede consultarse en <http://clic.ub.edu/corpus/en/ancora>. [Última consulta el 31 de marzo de 2023].

<sup>36</sup> Disponible para consultas en <http://www.elgrial.cl/>. [Última consulta el 31 de marzo de 2023].

<sup>37</sup> Disponible para consulta en [http://www.corpus.unam.mx/rst/index\\_es.html](http://www.corpus.unam.mx/rst/index_es.html). [Última consulta el 3 de mayo de 2022].

<sup>38</sup> Disponible para consulta en <http://www.llf.uam.es/~sandoval/UAMTreebank.html>. [Última consulta el 31 de marzo de 2023].

Autónoma de Madrid; el *IULA Treebank*<sup>39</sup> (Marimon, Fisas, Bel, Vivaldi, Torner, Lorente, ... y Villegas, 2012; Rojo, 2016) y el *LSP Spanish treebank*<sup>40</sup>, desarrollados en el IULA de la Universidad Pompeu Fabra (Marimon, Bel, Fisas, Arias, Vázquez, Vivaldi, J., Morell y Lorente, 2014; Marimon y Bel, 2015; Rojo, 2016) y que cuentan con anotaciones sintácticas únicamente.

Por su lado, *Araknion-es*<sup>41</sup> (del grupo CLIC), contiene análisis sintácticos con árboles de dependencias (Rojo, 2016) y es parte del *Proyecto Araknion*, que tiene como objetivo «dotar al español y catalán de una infraestructura básica de recursos lingüísticos para el procesamiento semántico de corpus en el marco de la Web 2.0» (Proyecto Araknion, 2010).

Para el español costarricense, se creó el *treebank* del español *IPROCOLDI*, parte del *Corpus de Mensajes Presidenciales Costarricenses (CODIMEP-CR)*, de la Universidad de Costa Rica (Jara Murillo, 2013). Por medio de *FIPS (LATL)*, Jara Murillo postuló el formalismo *IML* (Iprocoldi Marked Language) para la anotación sintáctica, e incorporó el etiquetado de *EAGLES* (Expert Advisory Group on Language) en la anotación léxica y morfológica; con el *PoS tagger* de *FreeLing* efectuó dos niveles de análisis: el morfológico (o nivel de etiquetado) y el sintáctico (o nivel sintácticamente analizado). No se incluyeron estudios del nivel semántico.

Muchos otros corpus escritos en español han sido anotados lingüísticamente en relación con análisis morfosintácticos y sintácticos, como los presentados en diversos estudios publicados (ver en Moreno-Sandoval y Campillos-Llanos, 2013; Koza Orellana, 2015; Sancho Escrivá, Fanjul Peyró, de la Iglesia Vayá, Montell y Escartí Fabra, 2020; Marimon y Bel, 2015; Sandí, Marín y Casasola, 2016; Obando, Pulido y Gómez, 2020). También se han hecho estudios con corpus para el análisis de sentimientos (Zafra, Gómez y Navarro-Colorado, 2017); análisis de paráfrasis (Vila, Bertran, Martí y Rodríguez, 2015); para la clasificación de frases clave (Medina y Turmo, 2018); para la similitud semántica textual o *STS* (Wang, Castellón y Comelles, 2020); para la traducción automática (Oliver y Climent, 2011); o para otros elementos más específicos, como los pronombres reflexivos (Degraeuwe y Goethals, 2020).

---

<sup>39</sup>Disponible para consulta en <http://linghub.org/metashare/def4ca5a92c211e28763000c291ecfc85641ed789c0740bd9d4a677dcda861d1>. [Última consulta el 3 de mayo de 2022].

<sup>40</sup> Disponible para consulta en [http://lod.iula.upf.edu/resources/metadata\\_IULA-Spanish\\_treebank](http://lod.iula.upf.edu/resources/metadata_IULA-Spanish_treebank). [Última consulta el 31 de marzo de 2023].

<sup>41</sup> Disponible para consulta en <http://clic.ub.edu/corpus/en/araknion>. [Última consulta el 31 de marzo de 2023].

### 1.6.5 Antecedentes del corpus en español con anotaciones semánticas

A pesar de las muchas anotaciones de corpus en español mencionadas en párrafos anteriores, la anotación semántica ha sido el tema menos tratado para el español, aunque se ha comenzado a trabajar con varios *treebanks* y con corpus de mayor tamaño, como el *SenSem*<sup>42</sup> (Vázquez, Alonso, Capilla, Castellón y Fernández, 2006; Vaamonde, 2008; Castellón, Climent, Coll-Florit, Lloberes y Rigau, 2012; Gil-Vallejo, Castellón, Coll-Florit, Turmo, 2015); el *AnCora* (Taulé, Martí y Recasens, 2008); el *Iarg-AnCora*<sup>43</sup> (Peris y Taulé, 2010; Taulé, Peris y Rodríguez, 2016); el *ADESSE* (Vaamonde, 2008); el *CESS-ECE*<sup>44</sup> (Martí Antonín, Taulé Delor, Màrquez y Bertran, 2007); el *CoNLL 09* (Robles, 2010; Robles, 2011); el *Tibidabo*<sup>45</sup> (Marimon, 2015); y otros de menor tamaño creados *ad hoc* para su estudio semántico (García Marchena, 2009; Reyes, Montes, González y Pinto, 2013; López Rodríguez, 2020).

Las anotaciones lingüísticas en español, sean morfológicas, sintácticas o semánticas, se han dado entre procesos manuales, semiautomáticos y automáticos (ver 1.7.1.1), con una marcada tendencia hacia la automatización de los análisis lingüísticos con miras hacia el aprendizaje automático (Calvo y Gelbukh, 2008; García Marchena, 2009; Robles, 2010; Robles, 2011; Sandí Salazar, Marín Raventós y Casasola Murillo, 2016; Marín Raventós y Sandí Salazar, 2017) y con él, hacia la disminución de costos y de tiempo en el procesamiento del lenguaje natural (Navarro-Colorado, 2007).

Si bien la calidad de las etiquetas lingüísticas para el procesamiento del lenguaje natural ha progresado con el pasar de los años, los analizadores automáticos para el etiquetado semántico en español contienen numerosos problemas que requieren la intervención de un lingüista para que corrobore y corrija los resultados que se obtienen y, de ser así, alcanzar una automatización más fiable (Vaamonde, 2008, Ballesteros et al., 2016; Gerdes y Kahane, 2016; Colic y Rinaldi, 2019; Nivre et al., 2020; Pyvsaalo et al., 2015; Üstün et al., 2020; Reiss, 2023).

---

<sup>42</sup> Disponible bajo suscripción con pago en <https://catalog.ldc.upenn.edu/LDC2015T02>. [Última consulta el 31 de marzo de 2023].

<sup>43</sup> Disponible para consultas en <http://clic.ub.edu/corpus/en/iarg-ancora-en>. [Última consulta el 31 de marzo de 2023].

<sup>44</sup> Ver más información en la página web [http://lod.iula.upf.edu/resources/project\\_CESS-ECE](http://lod.iula.upf.edu/resources/project_CESS-ECE). [Última consulta el 31 de marzo de 2023].

<sup>45</sup> Ver más información en la página web [http://lod.iula.upf.edu/resources/metadata\\_TRL\\_Tibidabo\\_LSP\\_treebank\\_ES](http://lod.iula.upf.edu/resources/metadata_TRL_Tibidabo_LSP_treebank_ES). [Última consulta el 31 de marzo de 2023].

### 1.6.6 Modelos grandes de lenguaje

En el 2018, NLP tuvo su primer gran éxito en el aprendizaje de red neuronal auto supervisado a gran escala (Manning, 2022) y la arquitectura de transformadores, con la que se desarrollaron dos tipos de modelos de lenguaje de gran tamaño o LLM (*Large Language Model*): el modelo *BERT* y el modelo *Generative Pretrained Transformer (GPT)* (Manning, 2022; De Angelis, Baglivo, Arzilli, Privitera, Ferragina, Tozzi y Rizzo, 2023). Ambos modelos emplean un enfoque semisupervisado, con preentrenamiento no supervisado, y *CHATGPT*, además, tiene afinamiento supervisado (*supervised fine-tuning*), con el que aplican representaciones de lenguaje preentrenadas (Schulman, Zoph, Kim, Hilton, Menick, Weng, Ceron Uribe, Fedus, Metz, Pokorny, Gontijo Lopes, Zhao, Vijayvergiya, Sigler, Perelman, Voss, Heaton, Parish, Cummings, Nayak, ... Balcom, 2022; De Angelis, et.al, 2023).

Los modelos grandes de lenguaje (*Large Language Models*, LLM) son modelos fundacionales que se utilizan en tareas de NLP y de NLG (*Natural Language Generation*) (Dilmegami, 2023). Su construcción se basa en la arquitectura de los transformadores generativos preentrenados (*Generative Pre-trained Transformers*) o modelos de aprendizaje profundo (*Deep Learning*, DL), el cual es un tipo de ML que utiliza múltiples capas de procesamiento para descubrir patrones y estructuras en conjuntos de datos de grandes tamaños (Rusk, 2016), que generan textos similares a los producidos por humanos (Wilcox, Futrell y Levy, 2022; Awan, 2023; Lee, 2023; Dilmegani, 2023). Los transformadores emplean algoritmos de aprendizaje profundo para reconocer, resumir, traducir, predecir y generar contenido basado en el aprendizaje obtenido de conjuntos de datos masivos (Lee, 2023; Arwan, 2023; Dilmegani, 2023). Se preentrenan sin instrucciones específicas (aprendizaje no supervisado) usando volúmenes gigantescos de datos que han sido escritos en internet durante un periodo de tiempo extenso (Lee, 2023).

El conocimiento o la pericia de los transformadores se puede afinar dándole al modelo pequeñas cantidades de datos, como corpus anotados, para que se enfoquen en ellos y aprendan para una aplicación específica (*prompt-tuning*) (Lee, 2023; Dilmegani, 2023). La puesta a punto de los transformadores es supervisada y los métodos más comunes para entrenamiento de los modelos preentrenados son el afinamiento (*fine-tuning*) y por comandos (*prompt-tuning*), y este último puede ser por medio de aprendizaje en contexto (*in-context learning*) o por aprendizaje en cero-uno-pocos tiros (*zero-/one-/few-shot learning*) (Lee, 2023; Dilmegani, 2023; Pourpanah,

Abdar, Luo, Zhou, Wang, Lim, . . . Wu, 2023), que rigen de mayor a menor intervención en el entrenamiento de las redes neuronales del transformador (Fichtel, Kalo y Balke, 2021).

Las redes neuronales son un tipo de ML que fue diseñado para imitar la estructura del cerebro humano, con nodos interconectados que procesan la información que reciben por medio de valores de peso y estos valores determinan la fuerza de cada conexión entre los nodos (Deng y Lin, 2022): «cada *token* de entrada *i* está conectado a todos los demás tokens en la siguiente capa, con los pesos relativos de la conexión correspondientes a su importancia para predecir el *token* en la siguiente capa» (Wilcox, Futrell y Levy, 2022, pág. 13).

Desde la perspectiva del aprendizaje profundo, Devlin y Chang (2018) y Devlin, Chang, Lee y Toutanova (2019) presentaron un nuevo modelo de representación del lenguaje, llamado *BERT* (*Bidirectional Encoder Representations from Transformers*) o Representación de Codificador Bidireccional de Transformadores, que ofrece resultados de última generación en diferentes tareas de NLP, incluida la respuesta a preguntas (*question answering*) (Horev, 2018). *BERT* es un transformador que emplea técnicas de preentrenamiento del NLP basadas en redes neuronales desarrollado por Google (Martínez, 2019) y que necesita grandes cantidades de datos con anotaciones para su entrenamiento (Horev, 2018).

Continuando con Horev (2018), *BERT* es un modelo bidireccional que aplica el entrenamiento bidireccional de transformadores al modelado de lenguaje. La novedad de *BERT* es que este modelo tiene un sentido más profundo del contexto y el flujo del lenguaje que los modelos anteriores de una sola dirección. *BERT* aprende las relaciones contextuales entre las palabras en un texto, porque es capaz de leer toda la secuencia de las palabras de entrada de texto (*input*) a la vez, en lugar de leerlas de izquierda a derecha o de derecha a izquierda (Horev, 2018). Aunque se le considere un modelo bidireccional, Horev (2018) lo llama un modelo no direccional.

A partir de *BERT*, Canete, Chaperon, Fuentes, Ho, Kang y Pérez (2020) crearon *BETO*, un modelo entrenado con un corpus extenso en español. El corpus lo compilaron utilizando todos los datos en español de tanto Wikipedia como de todas las fuentes del *OPUS Project*. Los autores afirman que el tamaño del corpus era comparable con el empleado en la versión original de *BERT* (Canete et al. 2020). El entrenamiento de *BETO* se llevó a cabo con algoritmos computacionales automatizados (Departamento de Ciencias de la Computación Universidad de Chile, 2022).

Canete et al. (2020) emplearon la técnica de *Dynamic Masking* o enmascaramiento dinámico para el entrenamiento, la cual se refiere al uso de diferentes máscaras para la misma frase

en el corpus<sup>46</sup>. También consideraron *Whole-Word Masking* (WWM), empleada en la versión actualizada de *BERT*. Para evaluar el entrenamiento automático, Canete et al. (2020) convirtieron la versión original del *English GLUE benchmark (General Language Understanding Evaluation)* en las del punto de referencia de *GLUES* para el español.

Las tareas de NLP que Canete et al. (2020) evaluaron fueron: la inferencia del lenguaje natural (*Natural Language Inference*), entrenaron el modelo con una traducción automática del conjunto de datos *MNLI* y emplearon el set de desarrollo y de prueba del corpus *XNLI*. Los resultados los califican por medio de la exactitud (*accuracy*). Otra tarea fue el parafraseo, con la que aplicaron la versión multilingüe del conjunto de datos *PAWS-X* para determinar si dos frases son semánticamente equivalentes o no. Esta tarea se mide con la exactitud también.

Para el reconocimiento del nombre de entidades (*Name Entity Recognition* o *NER*), Canete et al. (2020) usaron el conjunto de datos de Tjong Kim Sang (2002) (citado en Canete et al. (2020)) y lo evaluaron con *F1 score*<sup>47</sup>. En cuanto a las etiquetas de POS, consideraron el subconjunto de *Universal Dependencies (v1.4) Treebank* y lo valoraron con la exactitud de las etiquetas POS predichas. Para la clasificación de documentos, emplearon el *Multilingual Document Classification Corpus* (MLDoc) del Reuters corpus y midieron su exactitud. Por medio de *Unlabeled Attachment Score* (UAS) y *Labeled Attachment Score* (LAS)<sup>48</sup> evaluaron el análisis de dependencias e implementaron el *Universal Dependencies v2.2 Treebank*. Finalmente, Canete et al. (2020) usaron tres versiones traducidas del Conjunto de datos estándar de respuesta a preguntas *SQuAD (Standard Question Answering Dataset)* para trabajar *Question Answering*: MLQA, TAR y XQuAD. Consideraron los sets de entrenamiento de MLQA y TAR, y los sets de prueba de MLQA y XQuAD. Esta tarea la evaluaron con *F1* y con coincidencia exacta.

Canete et al. (2020) obtuvieron muy buenos resultados en las tareas de NLP en comparación con el desempeño del *mBERT*, un modelo multilingüe de *BERT*. No obstante, cabe destacar que el entrenamiento de este modelo se basa en aprendizaje automático y no en procesos manuales con la supervisión directa de equipos lingüistas en la anotación de los corpus implementados.

---

<sup>46</sup> Más detalles sobre modelos de enmascaramiento en Sánchez-Junquera, Rosso, Montes-y-Gómez y Chulvi (2021).

<sup>47</sup> *F1 score* es una métrica que se aplica en modelos de clasificación en ML (Korstanje, 2021).

<sup>48</sup> UAS y LAS son métricas para medir las dependencias y las relaciones entre las palabras (Cothenet, 2019).

Un breve resumen de la historia reciente de los *Transformers* comienza con *BERT* (Devlin y Chang, 2018; Devlin, Chang, Lee y Toutanova, 2019), el cual no puede generar textos por medio de comandos (*prompts*) y ha sido usado por Google para interpretar búsquedas de lenguaje natural (Awan, 2023). Luego en 2018, OpenAI presentó *GPT-1* como prueba del concepto del modelo (Radford, Narasimhan, Salimans y Sutskever, 2018), aunque no tuvo acceso para el público en general. Este tipo de modelo generaliza el lenguaje humano basándose en el principio de metaaprendizaje con redes neuronales (De Angelis, et al, 2023), el cual le provee un set de destrezas y habilidades para reconocer patrones, con el que pueden predecir el siguiente elemento en el texto (*token*) o lo que se conoce como probabilidad contextual (Deng y Lin, 2022; De Angelis, et.al, 2023; Eloundou, Manning, Mishkin y Rock, 2023).

En 2019 OpenAI presentó *GPT-2*, pero solo estuvo disponible para la comunidad de ML y presentaba muchas limitaciones (Radford, Wu, Child, Luan, Amodei y Sutskever, 2019). En 2020 salió *GPT-3*, el cual superaba en entrenamiento a su predecesor *GPT-2* (Brown, Mann, Ryder, Subbiah, Kaplan, Dhariwal, ... y Amodei, 2020). Para el 2022, se dio a conocer *GPT-3.5* con *ChatGPT*, un modelo mejorado y enfocado en conversación (Schulman et al., 2022), al que cualquier persona puede tener acceso gratuito. Al igual que *BERT* y *BETO*, *ChatGPT* se entrenó con grandes cantidades de datos textuales tomados de la Internet, razón por la cual este sistema de interfaz hombre-máquina ha llegado a expresar comentarios inintencionados que se apartan de los valores del usuario con el que conversa, porque toma como válida tanto información verdadera como aquella que ha sido publicada sin referencias confiables (Schulman et al., 2022; OpenAI, 2023; De Angelis, et.al, 2023).

La versión más actualizada al momento de redactar este apartado es *GPT-4*, que requiere suscripción pagada, cuenta con menos errores (“*hallucinations*”)<sup>49</sup>, y su comportamiento puede cambiarse para que se ajuste más a los valores del usuario (Awan, 2023). Además de haber mejorado la alineación del modelo con respecto a incoherencias de las versiones anteriores, *GPT-4* es capaz de analizar entradas textuales o visuales y ha demostrado tener capacidades cada vez más avanzadas y un rendimiento de nivel humano en algunos exámenes profesionales y académicos, en aspectos como razonamiento del sentido común y traducción (Awan, 2023; Lee,

---

<sup>49</sup> Las mal llamadas alucinaciones de ChatGPT son, en realidad, contenidos generados por los LLMs que contradicen el conocimiento fáctico (Cao, Yang y Zhao, 2023), porque son inventados por el modelo, basándose en los patrones que ha aprendido. Estos errores son información inventada que no es verificada (Li, 2023) y que puede desviarse de la realidad del usuario.

2023). En el estado de la cuestión se presentarán estudios evaluando el desempeño de *ChatGPT* como anotador de tareas en NLP y NLG.

Es importante aclarar que en esta investigación no se detallará a profundidad este tipo de modelos ni su entrenamiento, puesto que se escapa a los límites establecidos para el desarrollo de los objetivos. No obstante, ha sido necesario introducir en este apartado algunos de los avances más destacados en NLP que se han conseguido durante el curso de la presente tesis. OpenAI ha desatado polémica con el lanzamiento público de *ChatGPT*, esto por las habilidades que este sistema de pregunta-respuesta tiene para crear conversaciones y documentos que se asemejan, para muchos, el comportamiento y el razonamiento del cerebro humano (Deng y Lin, 2022; Haleem, Javaid y Singh, 2023; Aljanabi, Ghazi, Ali y Abed, 2023; Chomsky, Roberts y Watumull, 2023, 8 de marzo; Lee, 2023).

### ***1.6.7 Estado de la cuestión***

De los estudios sobre corpus anotados en español revisados para esta tesis, varios de ellos fueron realizados semiautomáticamente, como en Vázquez, Alonso, Capilla, Castellón y Fernández (2006), que utilizaron un proceso automático para etiquetar los sentidos verbales del corpus *SenSem*, el cual cuenta con 25 000 oraciones y 750 000 palabras en español para los 250 verbos más usados en textos del ámbito periodístico. Vázquez et al. (2006) anotaron semánticamente los constituyentes relacionados con el predicado verbal únicamente, con la idea de no excluir el análisis de las frases subordinadas.

Vázquez et al. (2006) trabajaron el nivel léxico para el sentido verbal, el de los constituyentes para asignar la categoría sintagmática, la función sintáctica y el tipo de relación argumental con el verbo (argumento/adjunto) a los participantes de la oración y el oracional para caracterizar el significado del conjunto en anticausativa, antiagentiva, impersonal, reflexiva o recíproca. La anotación fue manual, en cuanto a que los anotadores tomaron primero un patrón prototípico de roles semánticos, luego delimitaron los constituyentes y, finalmente, asignaron los roles semánticos y la semántica oracional para cada ejemplo analizado.

El etiquetado manual les permitió a Vázquez et al. (2006) crear la anotación automática de la categoría y la función más frecuentes, por medio del sistema que las preselecciona. Esta anotación automática es validada o modificada por los anotadores, manualmente. La interfaz

utilizada para las anotaciones fue el *GRIAL*, de donde se puede consultar dicho corpus en línea<sup>50</sup>. Para esta investigación cabe señalar que, a pesar de que los objetivos que describen los autores son «describir y representar el comportamiento léxico, sintáctico y semántico de los verbos del español» (pág. 113), no detallan la base teórica o gramática que justifica la selección de los roles temáticos, sino que emplearon un patrón prototípico de roles semánticos, difícil de seguir o equiparar con aquellos empleados en las teorías más prominentes<sup>51</sup>.

Castellón, Climent, Coll-Florit, Lloberes y Rigau (2012) trabajaron también el corpus *SenSem*, aunque desde otra perspectiva, en donde anotaron semánticamente los núcleos nominales de los argumentos, tomando como base para los sentidos léxicos el *ESPWN1.6* y el *Multilingual Central Repository* (MCR), por medio del recurso de anotación léxico-semántica *WordNet 1.6*. La anotación morfosintáctica del corpus la realizaron automáticamente por medio de *FreeLing* (Lexical Computing, 2020), con la que identificaron qué sustantivos debían anotar semánticamente. Sin embargo, con esta herramienta tuvieron problemas para etiquetar apropiadamente las unidades léxicas multipalabra, así como también la discrepancia entre las lematizaciones de *FreeLing* y de *WordNet* (Princeton University, 2010), e irregularidades estructurales en *WN1.6* en relación con autohiponimia y falsa hiponimia.

Castellón et al. (2012) hallaron también problemas de carácter interlingüístico, como la falta de equivalencia en español para ciertos términos, desdoblamiento indebido de *synset* a causa de diferencias morfológicas y de sesgo cultural, puesto que el *WN1.6* partía de una traducción del *WN1.6* del inglés (Atserias, Climent, Farreres, Rigau y Rodríguez, 2000, citado por Castellón et al., 2012). El objetivo de Castellón et al. (2012) era desambiguar semánticamente los núcleos argumentales de *SenSem* para presentar «un corpus etiquetado sintáctico y semánticamente» (p. 35) y, a su vez, criticar analíticamente *WordNet 1.6* en español.

Con su estudio, además del etiquetado morfosintáctico, Castellón et al. (2012) implementaron una interfaz de anotación para “búsquedas por lema, categoría morfosintáctica e identificador de oración” (p. 35). Asimismo, establecieron operadores para identificar metáforas, metonimias, multipalabras y partitivos. Castellón et al. (2012) no detallaron tampoco la base

---

<sup>50</sup> Ver más en <http://grial.edu.es/sensem/corpus?idioma=ca>. [Última consulta el 16 de junio de 2023].

<sup>51</sup> Las anotaciones de este corpus se pueden consultar en <http://grial.edu.es/sensem/lexico?idioma=ca>. [Última consulta el 16 de junio de 2023].

teórica empleada en su desarrollo y dejaron de lado las explicaciones lingüísticas teóricas que justifican la elección de las etiquetas empleadas para los roles semánticos asignados en su modelo.

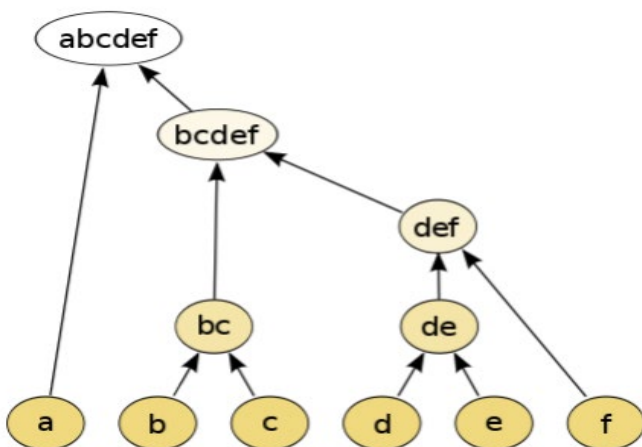
Gil-Vallejo, Castellón, Coll-Florit y Turmo (2015) también trabajaron con 20 sentidos verbales tomados del corpus *SenSem*. Para su investigación utilizaron un único sentido por verbo, con la idea de obtener modelos más precisos para la automatización de la clasificación verbal para el español. Los 20 sentidos verbales escogidos presentan diferentes esquemas sintácticos que pertenecen a varios campos semánticos y poseen diversos tipos de *iniciadores de eventos*, como causativos, agentivos y experimentadores.

Los siguientes son los sentidos verbales que Gil-Vallejo et al. (2015, pp. 43-44) emplearon en su anotación: para ‘estado’: ‘parecer’, ‘valer’ y ‘estar’; ‘comunicación’: ‘valorar’ y ‘explicar’; ‘cognición’: ‘gustar’ y ‘pensar’; ‘movimiento’: ‘perseguir’, ‘viajar’, ‘volver’ y ‘montar’; ‘cambio’: ‘abrir’, ‘cerrar’, ‘crecer’ y ‘morir’; ‘percepción’: ‘ver’ y ‘escuchar’; y para ‘actividad’ (social y corporal): ‘trabajar’, ‘dormir’ y ‘gestionar’.

A partir de los verbos seleccionados, Gil-Vallejo et al. (2015) construyeron un corpus *gold standard* basado en las estructuras sintácticas básicas descritas por Demonte y Bosque (1999) y en las construcciones de Levin (1993) (ambos textos citados en Gil-Vallejo et al., 2015). Asimismo, emplearon las estructuras sintácticas básicas de transitiva, intransitiva, ditransitiva, predicativa y atributiva. En cuanto a las construcciones, Gil-Vallejo et al. (2015) ofrecieron definiciones para la causativa prototípica, la anticausativa prototípica con ‘se’, la causativa de perífrasis, la anticausativa sin ‘se’, la voz media, la impersonal pronominal, el sujeto oblicuo, la reflexiva, la recíproca, la pasiva perifrástica, la pasiva refleja, el objeto cognado y la resultativa con ‘estar’. Para crear las clases verbales de este *gold standard* utilizaron un *clustering* o agrupamiento jerárquico aglomerativo, junto con los atributos y los sentidos verbales escogidos, como se aprecia en la Figura 2 (Gil-Vallejo et al., 2015, pág. 45).

**Figura 2**

*Modelización del clustering jerárquico aglomerativo*



En cuanto a los atributos lingüísticos empleados, Gil-Vallejo et al. (2015) listan los siguientes (págs. 45-46):

- Atributos semánticos de los argumentos:
  - Roles semánticos obtenidos mediante un mapeo jerárquico entre el *SenSem* y el *Lirics* (Bonial et al., 2011, citado por Gil-Vallejo et al., 2015).
- Incluyen 40 roles semánticos finos.
- Incluyen 16 roles semánticos abstractos.
- Incluyen 45 supersenses de *WordNet*, los cuales son categorías lexicográficas semánticas utilizadas para sustantivos, adjetivos o preposiciones (Jurasky y Martin, 2019).
- Incluyen 1000 términos de la ontología *SUMO*<sup>52</sup> que, al igual que los supersenses, los obtuvieron del núcleo de los argumentoso verbales de *SenSem*, los cuales están anotados con *synsets*<sup>53</sup>.
  - Atributos morfosintácticos, como función sintáctica, categoría morfológica y construcción (topicalización o destopicalización del sujeto lógico, reflexividad o impersonalidad).
- Aspecto oracional, como estado, evento, proceso.

<sup>52</sup> Más sobre *SUMO* en su página web <http://www.ontologyportal.org/>. [Última consulta el 31 de marzo de 2023].

<sup>53</sup> “Synset es un tipo especial de interfaz simple que está presente en NLTK para buscar palabras en WordNet. Las instancias de Synset son agrupaciones de palabras sinónimos que expresan el mismo concepto. Algunas de las palabras tienen solo un Synset y otras tienen varios” (Gupta, 28 de enero de 2019).

A partir de su investigación, Gil-Vallejo et al. (2015) concluyen que la combinación de *supersenses* y funciones sintácticas contribuye a generar una clasificación automática similar a la del *gold standard* en un mayor número de ocasiones, con lo que afirman que las funciones sintácticas combinadas con roles semánticos abstractos o con *supersenses* tienen un papel fundamental en la automatización de la clasificación verbal para el español.

El trabajo de Gil-Vallejo et al. (2015) confirma que es imprescindible incluir los análisis sintácticos a la hora de etiquetar roles semánticos en un corpus que, eventualmente, podría ser utilizado para el procesamiento del lenguaje natural y el aprendizaje automático. Sin embargo, no describen la lista de roles semánticos que ha sido empleada en la automatización de esta tarea, la cual podría haber sido contrastada con la que se trabaja en la presente tesis. La única manera de acceder al etiquetado de Gil-Vallejo et al. (2015) es a través de la plataforma *GRIAL* (Grup de Recerca Interuniversitari en Aplicacions Lingüístiques, s.f.), en donde se pueden hacer consultas por verbo, como se detallará en el apartado de la metodología y en el II capítulo de este proyecto.

Martí Antonín, Taulé Delor, Màrquez i Villodre, y Bertran (2007) presentaron un trabajo similar al de Vázquez et al. (2006), en el que exponen la metodología para la construcción de un *treebank* multilingüe en catalán, castellano y euskera, llamado *CESS-ECE*. Con un enfoque basado en corpus, anotaron automáticamente los corpus de las tres lenguas con 500.000 palabras cada uno, en tres diferentes niveles de descripción lingüística: el morfológico, con anotación automática; el sintáctico superficial, con anotación automática y el sintáctico profundo, con anotación manual (por medio *TreeTrans*); el semántico, con anotaciones manuales para los sentidos de nombres (por medio de *WordNet*) y anotaciones semiautomáticas para los papeles temáticos, en ambos corpus.

Martí Antonín et al. (2007) dividieron el corpus en dos subcorpus: en primer lugar, el *corpus-origen*, del que obtienen los léxicos verbales *CESS-LEX-ESP* (con 1.482 verbos para el español), y en el que se explicita la relación entre las funciones sintácticas y la estructura argumental, y la temática de cada predicado. Y, en segundo lugar, el corpus-prueba que utilizaron para comprobar el rango de cobertura y la calidad de la anotación semántica automática. Para cada sentido del verbo, se detallan *todos* los esquemas sintácticos en los que aparece. En el análisis sintáctico profundo utilizaron las funciones sintácticas de sujeto (SUJ), objeto directo (CD), objeto indirecto (CI), complemento de régimen (CREG), complemento agente (CAG), predicativo (CPRED), atributo (ATR), circunstanciales (CC, CCT, CCL). En la Figura 3, Martí Antonín et al. (2007) etiquetan el número de sentido (01); la clase semántica, la cual corresponde a la clase de

verbos transitivos agentivos (ELS4.2); la proyección sintáctico-semántica (SUJ/AGT, SUJ/PAT), y las alternancias de diátesis en las que puede ocurrir el verbo (voz ACTIVA/PASIVA) para el lema ‘valorar’ (Martí Antonín et al., 2007, p.70):

### Figura 3

*Léxico verbal para el verbo VALORAR*

```
valorar-01 synset: 00445269v
ELS4.2
SUJ Arg0##AGT
CD Arg1##PAT
CC ArgM##ADV
EJ: "CCOO valoró la reducción del paro..."

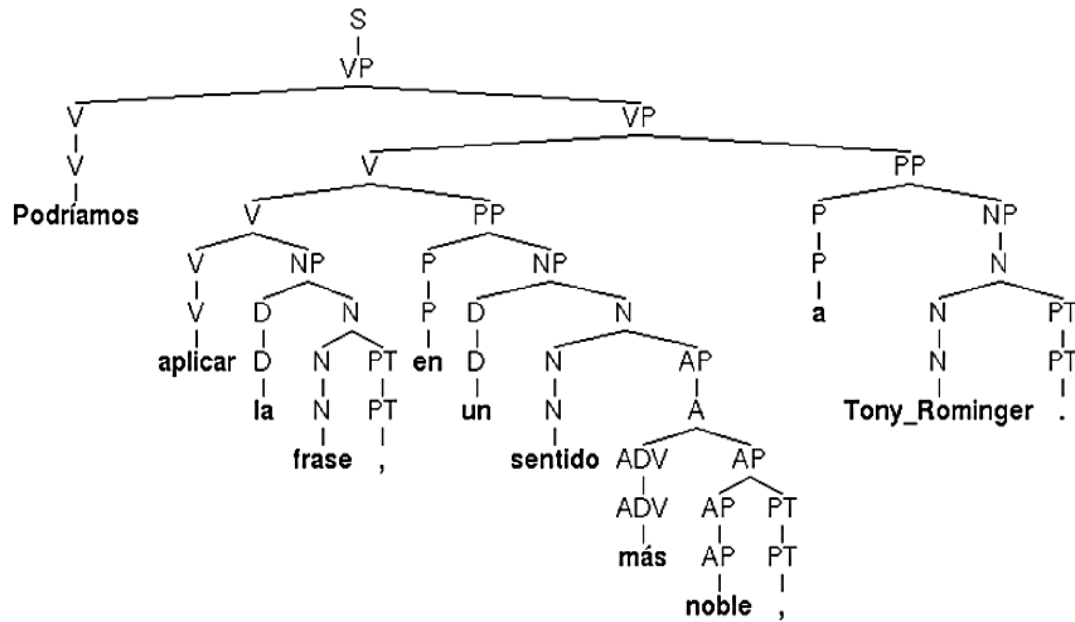
+PASIVA
SUJ Arg1##PAT
CAG Arg0#por#AGT
CC ArgM##ADV
EJ: "Las reservas en oro se valoran
    en_base_a 300_dólares estadounidenses..."
```

A pesar de la detallada descripción de las reglas para el procesamiento automático y de las etiquetas utilizadas, Martí Antonín et al. (2007) no definen una lista concreta de roles semánticos con la que anotaron el corpus, así como tampoco mencionan la teoría semántica o gramatical con la cual justificaron sus etiquetas, sino que solamente mencionan la proyección sintáctico-semántica con la que programaron la anotación automática.

Por su lado, Marimon (2015) parte de un set de 4000 oraciones en español tomadas de *AnCora* para anotar sintácticamente y semánticamente el corpus llamado *Tibidabo*. *AnCora* es un corpus que ya cuenta con anotaciones lingüísticas realizadas manualmente; sin embargo, Marimon (2015) ofrece un paradigma diferente, porque emplea una gramática léxica de análisis profundo basada en la Gramática Sintagmática Nuclear (*Head-driven phrase structure grammar*, HPSG) e implementada en el sistema *Linguistic Knowledge Builder* (LKB) (llamado *Spanish Resource Grammar* (SRG)) para crear anotaciones automáticas en lugar de anotaciones manuales humanas. Las anotaciones las ejecuta por medio del ambiente de *Deep Linguistic Processing with HPSG Initiative* (DELPH-IN). La mayor innovación de Marimon con *Tibidabo* tiene que ver con que hace la distinción entre argumentos y adjuntos, tanto en la estructura de dependencias como en la estructura semántica. Un ejemplo de su anotación se muestra en las Figuras 4 y 5:

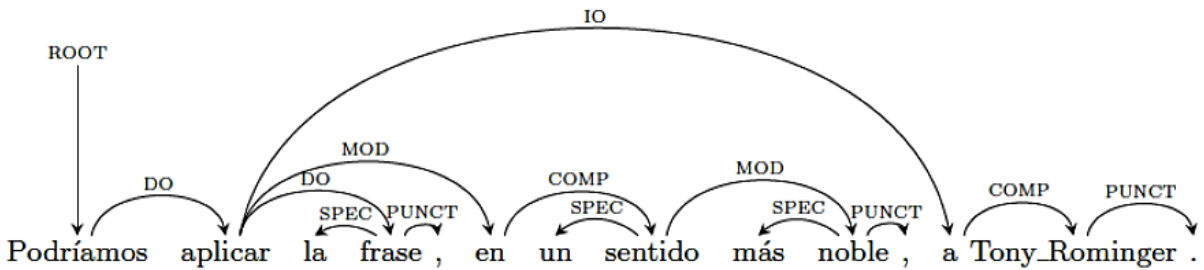
**Figura 4**

*Estructura de constituyentes*



**Figura 5**

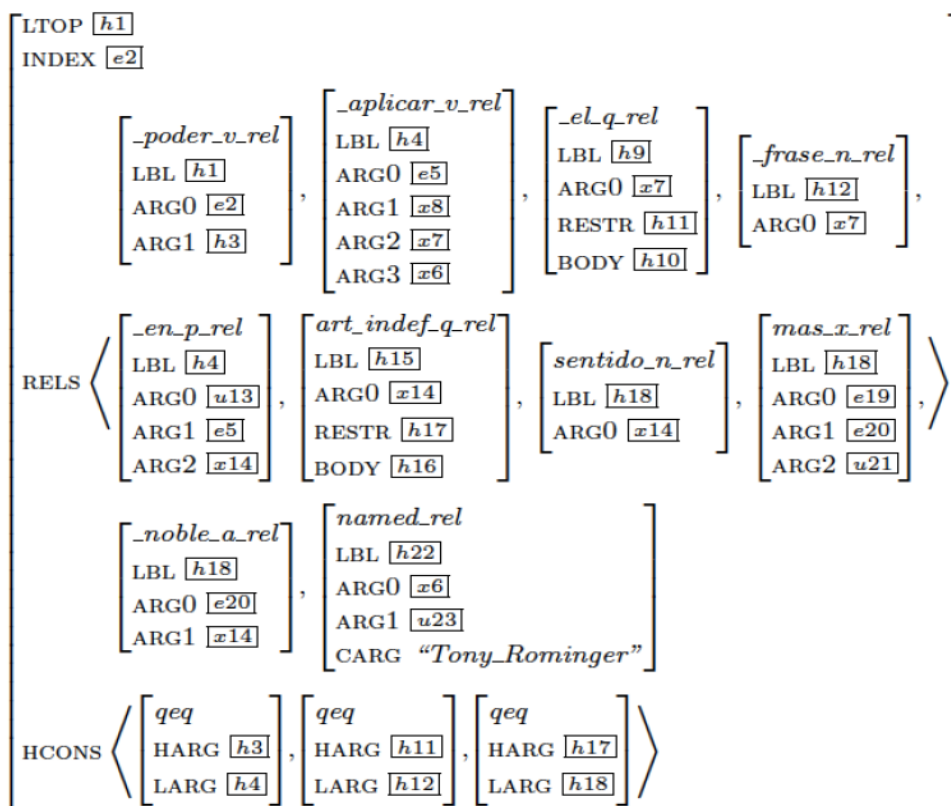
*Estructura de dependencias*



En la Figura 4 se presenta la estructura arbórea de los constituyentes de la oración. En la Figura 5, Marimon (2015) distingue entre argumento y adjunto dentro de los NP, AP, PP y ADVP, en donde los argumentos se etiquetan como COMP y los adjuntos como MOD.

**Figura 6**

*Estructura semántica*



En Figura 6, se muestra la distinción entre argumentos y adjuntos, puesto que los adjuntos: hacen que sus elementos modificados sean sus argumentos, mientras que estos estados se invierten para los argumentos: para el verbo aplicar ('aplicar') –x8, x7 y x6– se etiquetan como ARG1, ARG2 y ARG3 en la relación del verbo, mientras que la variable de evento del verbo –e5– identifica el ARG1 de la relación de la preposición que encabeza el PP que modifica el verbo [...]. Las etiquetas de dependencia semántica para los argumentos reciben un argumento numérico (es decir, ARG0, ARG1, ARG2, ARG3 y ARG4), además de una etiqueta de función temática.<sup>54</sup> (Marimon, 2015, párr. 13)

<sup>54</sup> Traducción propia del texto original: “Tibidabo also provides a clear distinction between arguments and adjuncts in the semantic structure, where adjuncts make their modifiees their arguments, whereas these statuses are reversed for arguments. This is illustrated in Figure 3, which shows the semantic structure of Example 1, where the arguments of the verb aplicar (‘to apply’) –x8, x7 and x6– are labelled as ARG1, ARG2 and ARG3 in the relation of the verb, whereas the verb’s event variable –e5– identifies the ARG1 of the preposition’s relation heading the PP that modifies the verb [...]. Here, semantic dependency labels for arguments are given a numerical argument (i.e., ARG0, ARG1, ARG2, ARG3 and ARG4) plus a thematic role label”.

La explicación sobre la anotación de *Tibidabo* por parte de Marimon (2015) es bastante difusa y su autora no incluye detalles sobre cuáles roles semánticos se etiquetaron. Sin embargo, cabe rescatar que la autora pretendía comparar sus resultados con aquellos obtenidos previamente con otras anotaciones semánticas ya realizadas en *AnCora*. Marimon (2015) varía el etiquetado, porque incluye las relaciones de dependencia entre los argumentos y los adjuntos con el verbo principal de la oración estudiada en la estructura sintáctica y en la semántica, un aspecto de gran importancia para la presente tesis, porque aquí se etiquetarán no solo los argumentos verbales, sino que también los adjuntos, de los que se destacarán los de lugar y de tiempo.

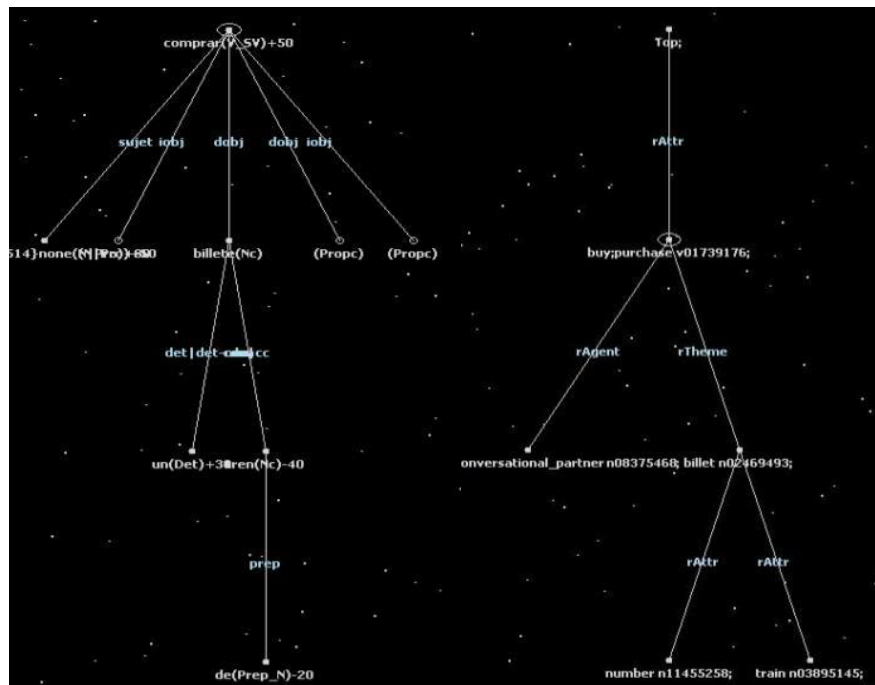
En cuanto a los sistemas de pregunta-respuesta, García Marchena (2009) presenta la arquitectura de los agentes de diálogo de *VirtuOz*, la cual emplea, como motor lingüístico, un analizador sintáctico-semántico que tiene como base la gramática formal de la Teoría Sentido-Texto, y una estructura de circuitos para la gestión de diálogos en español. García Marchena explica que, a diferencia de los modelos como *ELIZA*, considerados de primera generación, y los tipos *ALICE* y *PANDORA*, de segunda generación, los *chatbots* de esta empresa se inscriben dentro de los modelos de tercera generación, en cuanto a que integra un modelo lingüístico completo en su motor.

Los agentes de diálogo en *VirtuOz* realizan análisis en los niveles fonético, morfológico y sintáctico (tomando como base a Tesnière (1959), citado por García Marchena (2009)). Finalmente, lleva a cabo un análisis semántico con grafos semánticos para dar el significado de la palabra y su rol temático. Para este estudio, García Marchena (2009) explica que se requirieron 4 elementos básicos: una gramática de restricciones (Teoría Sentido-Texto), un léxico, una ontología en la que cada término léxico ocupa un lugar, y una lista de reglas de linealidad que asegura la estructura de dependencias más apropiada.

En el motor lingüístico, *VirtuOz* recurre al mecanismo de unificación para asegurar las asociaciones entre las diferentes fases en su implementación con el modelo llamado *GUST* (Gramática de Unificación polarizada Sentido-Texto). Este modelo de AI añade información en los nodos lingüísticos, de manera que determina qué categorías se combinan entre sí. *GUST* polariza los nodos presentes en el corpus etiquetado indicando si estos están saturados o no. Los nodos se conectan con arcos y estos arcos contienen la información sintáctica y semántica, indicando el rol semántico de esta relación, por medio de generalizaciones (ver la Figura 7 (García Marchena, 2009, págs. 7-8)).

**Figura 7**

*Representación sintáctica profunda y semántica*



Como metodología, García Marchena (2009) indica que emplearon un *tagger* con el algoritmo de Viterbi y las cadenas de Markov sobre el corpus semi manualmente etiquetado, con la idea de calcular las probabilidades de aparición de cierta categoría y las frecuencias de trigramas, esto último para determinar qué categoría adquiere una palabra nueva y para asignar la construcción sintagmática correspondiente. Este motor lingüístico también cuenta con un marcador de sinonimia que asigna automáticamente el significado a una palabra a partir de un sinónimo; un marcador para las colocaciones (lemas compuestos) y otro marcador para detectar usos vulgares de la lengua.

García Marchena (2009) comenta que la confección de un motor lingüístico para el español reveló ciertos problemas en el uso de un sistema de representación arborescente. Entre estas dificultades se detalla, principalmente, la baja precisión lingüística producto de las generalizaciones sistemáticas que se dan en la asociación entre las relaciones sintácticas y semánticas. Ejemplo de esto son los pronombres personales de sujeto que, cuando no se expresan sintácticamente (*pro-drop*), el grafo semántico resultante carece de la información semántica de sujeto.

Para solucionar esto, en *VirtuOz* emplearon la información de la morfología verbal de donde obtienen el rol semántico y así producen un grafo semántico completo, aunque García Marchena (2009) considera que, lingüísticamente, esta solución no es suficiente (Figura 7). Asimismo, otro asunto recurrente en español es el etiquetado y el análisis de los clíticos (Degraeuwe y Goethals, 2020), para lo cual implementaron un sistema de ponderaciones, dando prioridad a las estructuras más probables en español, empleando un sistema de subcategorización de complementos optativa para cada verbo.

Cabe subrayar que el uso de la Teoría Sentido-Texto como modelo lingüístico sacrifica la riqueza connotativa, porque se limita a reducir una expresión lingüística a su significado y deja de lado las excepciones que son comunes en la lengua española y que podrían estar presentes en el corpus compilado. Empero, según García Marchena (2009), la aplicabilidad computacional de este modelo lingüístico ha tenido éxito dentro de la AI, particularmente en su aplicación a los lenguajes técnicos, por lo que la convierte en una teoría lingüística con la que se obtienen resultados aceptables desde la perspectiva de la automatización.

Reyes, Montes, González y Pinto (2013) también toman como base teórica a Tesnière (1976) y a Halliday (1994) (citados por Reyes, et al., 2013) para responder qué rol semántico lleva a cabo cada entidad relacionada con un evento. Tesnière y Halliday consideran que los elementos fundamentales en una oración son los actores (actantes), la acción (verbo) y el decorado (circunstantes). Para esto, Reyes et al., (2013) eligen un corpus de cinco periódicos electrónicos mexicanos, entre marzo y noviembre de 2011. Del corpus extraen 3032 pares de sentencias, de las cuales 2274 fueron para entrenamiento, y 758 para pruebas. Su objetivo era clasificar automáticamente roles semánticos en sustantivos por medio de cuatro algoritmos diferentes: *K-Nearest Neighbors (K-NN)*, *Support Vector Machine (SVM)*, *Naive Bayes (NaïveBayes)* y *C4.5*; y determinar cuál de los algoritmos empleados ofrecía mejores resultados sobre el corpus trabajado.

En su investigación, Reyes et al., (2013) anotaron cinco categorías de los roles temáticos: Evento:Agente (EA); Evento:Objeto (EO); Evento:Beneficiario (EB); Evento:Circunstante (EC), para fuerza e instrumento únicamente; y Evento:Otro (EX). Para obtener características sintácticas y semánticas utilizaron la herramienta *FreeLing*, de las cuales las sintácticas fueron la posición de la entidad, la distancia de la entidad, la información morfológica del evento, la longitud del evento, la información morfológica del núcleo de la entidad, la entidad definida y la longitud de entidad. Y, en cuanto a las características semánticas, manejaron entidad nombrada, tipo de preposición,

tipo de entidad (nombre propio/común) e hiperónimo de la entidad. En adición a lo anterior, trabajaron con dos características contextuales: el contexto izquierdo y el contexto derecho, de las que se extrae la categoría gramatical.

Los autores concluyen que, de los cuatro algoritmos analizados, el *C4.5* es el clasificador con el que recibieron mejores resultados. Es importante destacar que este trabajo emplea una lista limitada de los posibles roles semánticos que se tienen en español; sin embargo, ofrece un aporte del comportamiento de algoritmos para el procesamiento automatizado de roles temáticos en español y para la evaluación de corpus anotados con este objetivo en otras investigaciones.

López Rodríguez (2020) aplica la Terminología Basada en Marcos para la clasificación de verbos en textos médicos, por medio de la metodología de LC y de la plataforma *Sketch Engine*. López Rodríguez (2020) elabora un inventario de verbos que concurren con sintagmas nominales que representan el concepto de signo y síntoma, importantes en un corpus de lenguaje especializado. El corpus se compuso con textos sobre medicina de diferentes niveles de especialización y, para ampliar la lista de verbos, se recurrió al corpus *EsTenTen11*, de *Sketch Engine*, y al buscador de *Google*. El etiquetado, y la recuperación de verbos prototípicos y de patrones lingüísticos se llevó a cabo de forma semiautomática a través de la plataforma *FreeLing*.

López Rodríguez (2020) asoció plantillas con información semántica y sintáctica sobre las predicaciones verbales que contienen signo y síntoma como argumentos, con la ayuda de un esquema de marcos y submarcos (*Frame Semantics*). En su investigación, López Rodríguez (2020) sigue la gramática de Demonte y Bosque (1999), al igual que Gil-Vallejo et al. (2015), así como también la Gramática Sistémico-Funcional de Halliday y Matthiessen (2004) y la Gramática Léxica de Faber y Mairal (1999) (citados en López Rodríguez, 2020). Esta investigadora asignó macrorroles y proto roles (proto agente / proto paciente) a los argumentos de los verbos estudiados, con sus funciones gramaticales y para los complementos circunstanciales asignó roles temáticos más específicos.

### **1.6.7.1 ChatGPT**

*ChatGPT* ha estado bajo riguroso estudio por distintas academias desde que se dio a conocer al público, y ejemplo de esto son los estudios de Qin et al. (2023), quienes analizaron tareas de razonamiento, inferencia de NL (*Natural Language*), diálogo, respuesta a preguntas (*question answering*), resumen, NER y análisis de sentimientos por medio de aprendizaje tiro cero

(*zero-shot learning*), de donde obtuvieron que el desempeño de *ChatGPT* fue superado por los modelos que habían sido entrenados con métodos *fine-tuning*. Anteriormente, en Zhang et al. (2022) *ChatGPT* había producido resultados de última en la detección de postura de los autores de textos.

Por su lado, Zhong et al. (2023) compararon *ChatGPT* con *BERT* y con *RoBERTa* en aprendizaje *fine-tuning* para las tareas de análisis de sentimientos, aceptabilidad lingüística, parafraseo, similitud textual, inferencia de NL y en respuesta a preguntas. En relación con *BERT*, los resultados fueron comparables; mientras que *RoBERTa* se desempeñó mejor que *ChatGPT*. En cuanto a traducción automática, Hendy, Abdelrehim, Sharaf, Raunak, Gabr, Matsushita, ... y Awadalla (2023) encontraron que los modelos *ChatGPT* logran una calidad de traducción muy competitiva para idiomas de altos recursos cuando se le compara con otros modelos completamente supervisados como *SOTA (state-of-the-art models)*, mientras que tienen capacidades limitadas para idiomas de pocos recursos.

Wang, Xie, Ding, Feng y Xia (2023) compararon el análisis de sentimientos de *GPT-3* con entrenamiento en modo *zero-shot*, con el de *BERT* y otros modelos *SOTA*, entrenados en modo *fine-tuning*. Estos investigadores encontraron un desempeño impresionante de *ChatGPT* en tareas de clasificación, en contraste con *BERT*, pero un poco atrás con respecto a modelos *SOTA*. Así también, *ChatGPT* tuvo menos precisión en tareas de extracción de información de sentimientos, pero una habilidad impresionante en la inferencia de sentimientos. *ChatGPT* obtuvo un desempeño comparable en tareas de extracción de los pares emoción-causea con respecto a *SOTA*. Para el fenómeno de cambio de polaridad, *ChatGPT* predijo con más precisión que *BERT*. Wang et al. (2023) concluyen que el modelo *ChatGPT* demostró su potente análisis de sentimientos de dominio abierto (*open-domain*), aunque todavía tiene limitaciones en otros dominios específicos.

Kuzman et al. (2023) comparan *ChatGPT* con *XLNet* en inglés y en eslovaco, y encuentran que *ChatGPT* supera a *XLNet*, principalmente en inglés, con mejores resultados en la identificación automática de géneros de textos (*News, Legal, Promotion*), una tarea que estos investigadores consideran difícil de realizar por humanos. Esto último porque las campañas de anotaciones manuales a menudo resultan en un pobre acuerdo entre anotadores y una fiabilidad escasa en los datos anotados, de aquí que Kuzman et al. (2023) lleguen a sugerir que *ChatGPT* podría sustituir este tipo de campañas humanas, si se mejoran los '*prompts*' y el tipo de corpus con el que se trabaje.

En los párrafos anteriores se expusieron algunos estudios en donde se ha comparado el comportamiento de modelos *GPT* con otros modelos de LLM en cuanto a tareas automatizadas de NLP. A continuación, se presenta el estado de la cuestión de investigaciones en donde el modelo se compara con anotaciones humanas.

Huang, Kwak y An (2023) encontraron que *ChatGPT* tiene potencial como categorizador de discursos de odio implícito en el conjunto de datos *LatenHatred*, porque el modelo identificó correctamente 80 % de los tuits que contenían dicho tipo de discurso a través de un comando simple (*prompt*). De acuerdo con los autores, las explicaciones dadas por *ChatGPT* superan las de los anotadores humanos y argumentan que los valores que obtuvieron en la confiabilidad de los datos anotados son suficientes para prescindir de anotadores humanos eventualmente. Sin embargo, Huang et al. (2023) advierten *ChatGPT* tiende a ser convincente con sus respuestas y que esto podría convertirse en un riesgo para las personas legas que hagan uso del modelo.

Al igual que Kuzman et al. (2023), Gilardi et al. (2023) opinan que *ChatGPT* es un anotador potencial en tareas de NLP, porque representa menores costos que los de un grupo de anotadores humanos (*crowd-workers*), como los de MTurk<sup>55</sup> o, incluso los de anotadores entrenados, como asistentes de investigaciones o estudiantes de posgrado. Gilardi et al. (2023) emplearon el entrenamiento *zero-shot* en *ChatGPT* para que este modelo etiquetara tuits en cinco tipos de tareas diferentes: relevancia, postura, temas y dos tipos de detección de fotogramas. Para lograr esto, emplearon los mismos *codebooks* o instrucciones que usaron con los anotadores de MTurk.

Gilardi et al. (2023) encontraron que en cuatro de las cinco tareas *ChatGPT* tuvo mayor precisión (72.8 % para relevancia y 78.7 % para postura o actitud) que los anotadores de MTurk con respecto al *gold standard* que había sido anotado por dos estudiantes de maestría en ciencias políticas, entrenados en estas tareas. Y, para el acuerdo entre anotadores, el de *ChatGPT* estuvo entre un 84 - 95 %, dependiendo de la temperatura o grado de aleatoriedad (0.2 y 1), mientras que el de los anotadores humanos entrenados estuvo entre 50 y 78.3 %, dependiendo de la tarea y de su dificultad. Con esto Gilardi et al. (2023) demuestran que, aún entre humanos entrenados, no es sencillo conseguir un acuerdo alto en este tipo de tareas.

En la misma línea, Zhu, Zhang, Haq, Hui y Tyson (2023) quisieron comprobar la utilidad de *ChatGPT* en la anotación de la postura, detección y análisis de sentimientos, discurso de odio

---

<sup>55</sup> Disponible en <https://www.mturk.com/>. [Última consulta el 25 de mayo de 2023].

y detección de *bots*, y comparar su desempeño con aquel llevado a cabo por grupos de humanos (*crowd-source*). Para esto emplearon un solo *prompt* para recrear la anotación de cinco conjuntos de datos que habían sido previamente anotados por personas. Zhu et al. (2023) encontraron que *ChatGPT* dio una precisión de 0.609, un alto desempeño en el análisis de sentimientos, un 64.9 % de tuits anotados correctamente y un puntaje F1 de 0.65. Los investigadores concluyen que con su estudio se abren posibilidades para explorar más el uso de *ChatGPT* como herramienta para tareas de NLP.

Similar a Gilardi et al. (2023), Törnberg (2023) confirmó con su estudio que *ChatGPT-4*, con entrenamiento *zero-shot* en temperaturas alta y baja, superó a un grupo de anotadores humanos de MTurk y a otro grupo de clasificadores expertos humanos cuando anotó tuits identificando afiliaciones políticas. Törnberg encontró que *ChatGPT-4* obtuvo mayor precisión y fiabilidad que los anotadores humanos, e igual o menor parcialidad (*bias*) que los clasificadores humanos. Törnberg concluye que *ChatGPT-4* tendrá un impacto sustancial en el uso de datos textuales en ciencias sociales, por cuanto permitirá llevar a cabo investigaciones interpretativas a gran escala. Törnberg agrega que el uso de LLMs como *ChatGPT-4* implicará bajos costos y replicabilidad en los experimentos, porque se pueden emplear las mismas instrucciones que las dadas a anotadores humanos.

De opinión opuesta a los hallazgos de Kuzman et al. (2023), Huang et al. (2023) y Gilardi et al. (2023), Reiss (2023) asevera que *ChatGPT*, al igual que los anotadores humanos, es de carácter no determinista, por lo que entradas idénticas pueden llevar a diferentes salidas y esto da como resultado poca consistencia en sus respuestas y, por lo tanto, lo convierte en una herramienta poco confiable en estudios académicos. Reiss (2023) comprobó la fiabilidad de *ChatGPT* con aprendizaje *zero-shot* cuando anota y clasifica textos de páginas web en noticias/no noticias (*news/no news*). Para esto, Reiss (2023) varió los parámetros del modelo en cuanto al grado de aleatoriedad (*temperature*) y los *prompts*, y también repitió las mismas entradas para verificar si *ChatGPT* ofrecía respuestas consistentes. El autor evaluó los resultados de la fiabilidad intracodificador con el coeficiente alfa de Krippendorff, esto es, comparando las respuestas de *ChatGPT* con las que había dado el mismo modelo con otros parámetros de temperatura, con los cambios en el *prompt* y con las repeticiones de las mismas preguntas.

Reiss (2023) reporta que cuando cambia solo la temperatura del modelo y repite el *prompt* 10 veces, resulta en  $\alpha = 0.75$ . Cuando cambia el *prompt*, pero mantiene las demás variables iguales,

obtiene  $\alpha = 0.60$ . Finalmente, cuando repite las mismas entradas del *prompt* con una temperatura menor,  $\alpha = 0.91$ ; mientras que, con una temperatura mayor,  $\alpha = 0.85$ . Cuantas más repeticiones efectuó Reiss, mejores valores de  $\alpha$  percibió. No obstante, este autor aconseja precaución cuando se emplee *ChatGPT* con aprendizaje *zero-shot*, y recomienda que los datos que se obtengan sean comparados con datos anotados por humanos, porque la variación de los comandos en este modelo causa respuestas inconsistentes que podrían tomarse por válidas. Pero, advierte, muchos investigadores no se sentirán atraídos a repetir *prompts* para etiquetar textos de gran tamaño a la hora de emplear un modelo automatizado, aunque este provea mejores resultados.

Finalmente, Chan, Cheng, Wang, Jiang, Fang, Liu y Song (2023) evaluaron el desempeño de *ChatGPT* identificando las relaciones entre oraciones y las relaciones discursivas. Chan et al. (2023) aplicaron tres *prompts* para cada tarea y también tres tipos de métodos de aprendizaje: *zero-shot*, *zero-shot engineering* e *in-context* en trece conjuntos de datos. Para el *zero-shot*, dieron al modelo preguntas de opción múltiple para que las solucionara. Para el *zero-shot engineering*, diseñaron manualmente un *prompt* para mejorar el rendimiento de *ChatGPT*, con etiquetas de dependencias para que el modelo predijera la clase del nivel superior y la clase del segundo nivel simultáneamente. Las dependencias de las etiquetas y los conectores seleccionados entrenaron al LLM para que entendiera el sentido de cada relación discursiva.

En cuanto al aprendizaje en contexto, Chan et al. (2023) seleccionaron manualmente ejemplos de entrada del conjunto de datos y reformularon el ejemplo referencia para que cada clase fuera distinguible y fácilmente comprensible para el modelo. Con su estudio, los autores hallaron que *ChatGPT* tiene un desempeño fuerte en la detección y el razonamiento de las relaciones causales, pero no es competente en identificar el orden temporal entre dos eventos; en cuanto a las relaciones discursivas, las implícitas son todavía un desafío por superar; y en las tareas de análisis del discurso, *ChatGPT* se desempeña pobremente en los casos en donde se requiere la comprensión de la conexión de dos oraciones en un diálogo antes de reconocer la relación discursiva.

En resumen, los antecedentes revisados para este proyecto ofrecen las siguientes conclusiones: en primer lugar, se establece que los análisis sintácticos son imprescindibles a la hora de etiquetar los roles temáticos (Gil-Vallejo et al., 2015; García Marchena, 2009; Martí Antonín et al. (2007); Vázquez et al., 2006; Castellón et al., 2012; Reyes et al., 2013; Marimon, 2015 y López Rodríguez, 2020), puesto que los análisis sintácticos permiten determinar las

relaciones argumentales que existen entre el núcleo verbal y los participantes de la oración y, a partir de estas relaciones, se determinan y establecen los roles temáticos.

En segundo lugar, los trabajos y aportes conjuntos de Vázquez et al. (2006), Castellón et al. (2015) y Gil-Vallejo et al. (2015) dieron lugar a la plataforma del *GRIAL*, una herramienta que puede ser empleada como guía de anotación de etiquetas para los análisis de las tres estructuras que proponen en LFG para corpus en español.

En tercer lugar, se determina que *FreeLing* es la plataforma de anotación más comúnmente empleada, a pesar de seguir presentando diversos problemas de etiquetado automático para frases en español, y que, en definitiva, cualquier anotación automática requiere de la corroboración lingüística manual antes de proceder al entrenamiento automático de sistemas de pregunta-respuesta (Castellón et al., 2012; Reyes et al., 2013; López Rodríguez, 2020).

En cuarto lugar, existen varias listas de roles temáticos en español y no una única aplicable a todos los corpus que se estudien, porque con cuáles o cuántos roles temáticos se trabaje dependerá de la teoría lingüística que se tome como base y, principalmente, dependerá de los propósitos propios de cada investigación (Vázquez et al., 2006; Martí Antonín et al., 2007; García Marchena, 2009; Castellón et al., 2012; Reyes et al., 2013; Gil-Vallejo et al., 2015; Marimon, 2015 y López Rodríguez, 2020). La anotación manual tenderá a considerar una lista más extensa de roles temáticos, mientras que las anotaciones automatizadas tienden a trabajar con proyecciones (Martí Antonín et al., 2007) o con listas reducidas de proto roles (Reyes et al., 2013).

Finalmente, en cuanto a la automatización del aprendizaje profundo en modelos bidireccionales como *BERT* y *BETO*, o mucho más avanzados de LLM como *ChatGPT*, se considera oportuno que desde la Lingüística se ofrezcan a otras academias corpus anotados en español como fuente de entrenamiento (para *fine-tuning*) para mejorar los resultados en las métricas de las tareas de NLP en esta lengua; porque, a pesar de que varios estudios sugieren que *ChatGPT* está capacitado para llevar a cabo las tareas de clasificación y etiquetado de textos y que se podría pensar en dejar de lado las anotaciones humanas (Huang et al., 2023; Kuzman et al., 2023; Gilardi et al., 2023; Törnberg, 2023), la constancia y fiabilidad de las respuestas que este LLM produce están en tela de juicio (Reiss, 2023).

De 1.1 a 1.6 se ha provisto al lector con el tema, los objetivos, el problema a tratar y el contexto del que se parte, así como también el estado de la cuestión hasta el momento de la

redacción de esta investigación. En el siguiente apartado se propone el marco teórico con el que se trabajará esta tesis.

## **1.7 Marco teórico**

En este apartado se desarrolla el marco teórico con el cual se construirá tanto el análisis sintáctico como el semántico para obtener las relaciones argumentales entre los participantes de la oración con el verbo. En primer lugar, se explicará las necesidades lingüísticas de los sistemas de pregunta-respuesta en 1.7.1; de qué se trata la anotación lingüística en 1.7.2 y los tipos de anotación en 1.7.2.1. Luego se detallarán los roles temáticos en 1.7.3 y las relaciones sintáctico-semánticas de los argumentos verbales desde la Gramática Léxico-Funcional en 1.7.4.

### ***1.7.1 Las necesidades lingüísticas de los sistemas de pregunta-respuesta***

Las computadoras no pueden pensar por sí mismas, sino que solamente pueden imitar al ser humano, tanto como este le pueda enseñar a hacerlo (Pretz, 2021). Esta limitación ha sido uno de los principales problemas en la rama de la AI que sigue sin encontrar una solución y que continúa en constante desarrollo y progreso. La AI tuvo sus inicios cuando varios académicos, motivados por los avances tecnológicos en computación durante la Segunda Guerra Mundial, comenzaron a crear los primeros intentos de traducciones automáticas en la década de 1940 y 1950 (Gallagher, Rafferty y Wu, 2004; Perrián Pascual, 2012).

Los programas originales, creados desde una perspectiva matemático-estadística, produjeron traducciones gramaticalmente aceptables, pero presentaban numerosos errores causados por problemas de ambigüedad léxica. Y era de esperarse, porque sus aproximaciones matemáticas, y no lingüísticas, tenían como base una concepción simplista de las diferencias entre las lenguas naturales, de manera que estos estudiosos se limitaban a creer que la sintaxis, los vocabularios y un analizador morfológico eran herramientas suficientes para traducir automáticamente textos entre diversas lenguas (Perrián Pascual, 2012).

La falta de fundamentos teóricos lingüísticos detrás del procesamiento del lenguaje natural constituyó el principal impedimento para conseguir una automatización eficiente en las primeras traducciones (Perrián Pascual, 2012). Los programas eran «engañosamente inteligentes», como los califica Perrián Pascual (2012, p. 15), porque, aunque aplicaban los algoritmos correctamente, no eran capaces de desempeñarse con un buen rendimiento en tareas que implicaran extracción y

uso de información en lenguas naturales, como era el caso de los errores por ambigüedad léxica en las traducciones automáticas.

Por su lado, en los sistemas de pregunta-respuesta, desde *ELIZA* (Weizenbaum, 1966) hasta *Siri*, *Alexa*, *Cortana* o *Xiaolce*, el lenguaje natural ha sido piedra de tropiezo para el procesamiento del lenguaje natural, a causa de que se debe enseñar al sistema informático a llevar a cabo tareas que el cerebro humano realiza de forma innata, como lo es comprender las preguntas que se le hacen y responderlas adecuadamente. Las etiquetas lingüísticas que se anotan en corpus para entrenamiento facilitan la enseñanza del lenguaje natural al sistema computacional; no obstante, para mejorar los resultados del ML, estas etiquetas deben ser supervisadas por la pericia de lingüistas (Periñán Pascual, 2012).

Desde la perspectiva lingüística, los *chatbots* requieren, como mínimo, de tres conceptos lingüísticos esenciales para su desempeño: la sintaxis, la semántica y la pragmática (Perdigão, 2020); y para que su desempeño sea realmente óptimo, es preciso que estos programas hayan implementado el procesamiento del lenguaje natural con los análisis léxico, sintáctico, semántico, discursivo y pragmático de corpus lingüísticos. De estos tipos de análisis y de anotación se describirán a continuación los pertinentes para la anotación sintáctica y la semántica para los roles temáticos, puesto que los otros tipos quedan fuera de los objetivos propuestos para esta tesis.

### ***1.7.2 Corpus y anotación lingüística***

En adición a las características referidas para los corpus en el apartado de los antecedentes, estos también se clasifican entre los que están codificados y anotados, y los que no lo están (Rojo, 2021). Pustejovsky y Stubbs (2013) denominan conjuntos de datos (*datasets*) o simplemente ‘corpus’ a los que no están anotados, y como corpus anotado (*annotated corpus*) a un único conjunto de datos anotado con la misma especificación. La diferencia entre ellos se basa en que los llamados “anotados” han sido enriquecidos con uno o varios niveles diferentes de descripción lingüística y se utilizan para entrenar algoritmos de ML.

En ML se aplican tres tipos de algoritmos: los de aprendizaje supervisado, los de aprendizaje no supervisado y los de aprendizaje semi-supervisado; en donde los primeros requieren corpus anotados para su entrenamiento, los segundos no necesitan corpus con etiquetas, mientras que el tercero es una mezcla de estos dos, en donde se combinan aprendizajes supervisados y no supervisados (Pustejovsky y Stubbs, 2013). De aquí que sea sumamente

necesaria la creación de corpus anotados, tipo *gold standard*, con la temática del COVID-19 para entrenar automáticamente sistemas de pregunta-respuesta.

Normalmente, los corpus anotados son parte de otros corpus mayores como algunos de los vistos en los antecedentes, de los que se selecciona un conjunto de textos y, de estos, se etiquetan distintas capas con la descripción lingüística correspondiente para llevar a cabo análisis lingüísticos o pueden ser creados *ad hoc* para investigaciones más precisas.

### 1.7.2.1 Tipos de anotación lingüística

Las anotaciones en los corpus pueden ser realizadas manual, automática o semiautomáticamente, tomando las manuales como las que ofrecen mayor precisión y calidad, dado que la intervención humana ofrece mayor fiabilidad en el éxito del ML. No obstante, es común que, para facilitar el proceso de la creación de corpus anotados, se empleen herramientas que asignen etiquetas automáticamente y que luego estas sean supervisadas por lingüistas para corroborar que no existan errores que puedan sesgar los resultados que se buscan con el ML.

Siguiendo a Hinze, Heese, Schlegel y Paschke (2019), las anotaciones son información adicional que se asigna a un documento y estas constan de dos partes: el *text atom* o texto seleccionado para anotarse, y el *annotation content* o contenido de la anotación. Las capas de la anotación lingüística varían de acuerdo con los fenómenos de la lengua que se requieran explicar, como las etiquetas del análisis sintáctico, el semántico, el morfológico, el fonológico, el fonético, el léxico, del discurso, el pragmático y el de la estructura del texto (Pustejovsky y Stubbs, 2013). Cuáles o cuántos niveles de análisis lingüísticos se efectúen dependerá de qué tan profunda sea la descripción que se le quiera dar a los textos del corpus y para qué propósitos se construya dicho corpus. Para esta investigación interesan particularmente las capas sintáctica y semántica para obtener la relación argumental de los participantes de la oración con el verbo (argumento, adjunto).

La anotación sintáctica incluye tres partes: el etiquetado gramatical o *part-of-speech tagging* (POS *tagging* o POST), la estructura de frase y la estructura de dependencias (Pustejovsky y Stubbs, 2013). El etiquetado gramatical consiste en asignar el marcador o los marcadores de la clase léxica correcta a todas las palabras en las oraciones presentes en el corpus que se vaya a trabajar. Pustejovsky y Stubbs (2013) enfatizan en que el POS es un paso crítico para poder continuar con los análisis lingüísticos subsecuentes, porque este proceso determina la categoría de cada palabra en el enunciado y por medio del POS, por ejemplo, se puede determinar si una palabra

funciona como un verbo o como un sustantivo dentro del corpus. Ejemplo de este conflicto se aprecia en (1)-(4)<sup>56</sup>, en donde las palabras homógrafas “amo” y “aro” están subrayadas para demostrar el efecto:

(1) Amo el olor de tu perfume.

<b>Amo</b>	<b>el</b>	<b>olor</b>	<b>de</b>	<b>tu</b>	<b>perfume</b>	.
amar	el	olor	de	tu	perfume	.
VMIP1S0	DA0MS0	NCMS000	SP	DP2CSS	NCMS000	Fp

(2) El amo acarició su patita.

<b>Su</b>	<b>amo</b>	<b>le</b>	<b>acarició</b>	<b>su</b>	<b>patita</b>	.
su	amo	le	acariciar	su	pata	.
DP3CSN	NCMS000	PP3CSD0	VMIS3S0	DP3CSN	NCFS00V	Fp

(3) El aro de la bicicleta rodaba sin problemas.

<b>El</b>	<b>aro</b>	<b>de</b>	<b>la</b>	<b>bicicleta</b>	<b>rodaba</b>	<b>sin</b>	<b>problemas</b>	.
el	aro	de	el	bicicleta	rodar	sin	problema	.
DA0MS0	NCMS000	SP	DA0FS0	NCFS000	VMII3S0	SP	NCMP000	Fp

(4) Yo aro la finca con mis bueyes.

<b>Yo</b>	<b>aro</b>	<b>la</b>	<b>finca</b>	<b>con</b>	<b>mis</b>	<b>bueyes</b>	.
yo	arar	el	finca	con	mi	buey	.
PP1CSN0	VMIP1S0	DA0FS0	NCFS000	SP	DP1CPS	NCMP000	Fp

En (1)-(4), las palabras han recibido la etiqueta correspondiente según la función gramatical en la oración a la que pertenecen. El *parser* de *Freeling* emplea etiquetas con letras y números para definir los marcadores que le corresponden a cada palabra y, por razones de espacio, solo se explicarán los valores dados en los ejemplos (3) y (4) para “aro”.

En (3), “aro” recibe la etiqueta “NCMS000”, con los atributos de la categoría, el tipo, el género, el número, el caso, el género semántico y el grado de cortesía (los atributos que corresponden al caso y al género semántico no se especifican en EAGLES, y se les asigna el valor 0), de manera que se obtiene “aro” como un ‘nombre común masculino en singular’. En (4), “aro” tiene la etiqueta “VMIP1S0”, en la cual se indican los atributos de la categoría, el tipo, el modo, el tiempo, la persona, el número y el género (para los participios únicamente), dando entonces “aro” como un ‘verbo principal en indicativo presente, en primera persona singular’. Con el uso

<sup>56</sup> Los ejemplos en el marco teórico no son parte del corpus con el que se trabajará en la anotación de los roles temáticos y los resultados de POS se tomaron de *Freeling*: <https://nlp.lsi.upc.edu/freeling/demo/demo.php>

del POS se determina que, aunque “aro” se encuentre dos veces en este mini corpus, este vocablo ejerce dos funciones diferentes y que estas dependen de la categoría que se le asigne en cada oración en que aparezca.

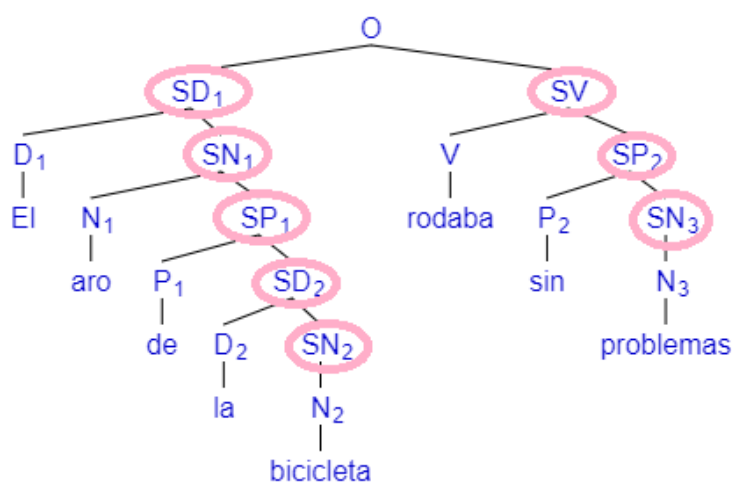
En adición a lo anterior, con las etiquetas del POS para las categorías gramaticales se crean las unidades sintácticas mayores o también llamadas sintagmas, las cuales son necesarias para establecer la estructura de las oraciones y la jerarquía entre los participantes. Los sintagmas son grupos de palabras que desempeñan una misma función dentro de los enunciados o, en palabras de Bosque y Gutiérrez-Rexach (2009), son los segmentos que:

son sensibles a distintos procesos gramaticales: sustitución por pronombres, desplazamientos, coordinación, elipsis, respuestas a preguntas. [...] También son los segmentos necesarios para explicar la estructura de otros más amplios, es decir, las unidades que los componen o las que permiten formarlos. (pág. 117)

Estos segmentos son el sintagma nominal (SN), el sintagma verbal (SV), el sintagma adjetival (SA), el sintagma preposicional (SP) y el sintagma adverbial (SAdv). Un breve ejemplo de sintagmas oracionales se presenta en la Figura 8, en donde los sintagmas de la oración (3) han sido resaltados en rosado<sup>57</sup>. Cabe aclarar que cada *parser* y cada investigador tienen sus propias siglas para designar las etiquetas de los sintagmas, por lo que para esta investigación y su etiquetado se emplearán las señaladas anteriormente en este párrafo.

**Figura 8**

*Árbol sintáctico de «El aro de la bicicleta rodaba sin problemas»*



<sup>57</sup> Árbol sintáctico elaborado con la herramienta *jsSyntaxTree* en <http://www.ironcreek.net/syntaxtree/>.

La estructura oracional se refiere a las secuencias específicas de palabras, las cuales introducen dos tipos importantes de relaciones entre las palabras: el orden en que aparecen (orden de precedencia) y la jerarquía o relaciones de dependencia entre las palabras. La Figura 8 muestra la jerarquía y el orden de precedencia de las palabras que conforman la oración del ejemplo (3), en donde los núcleos de los sintagmas dependen del sintagma mayor que está en una posición más alta en el árbol; de manera que SN<sub>3</sub> depende de SP<sub>2</sub> y este depende de SV y SV depende de O; mientras que SN<sub>2</sub> depende de SD<sub>2</sub>, este depende de SP<sub>1</sub> y SP<sub>1</sub> depende de SN<sub>1</sub> y este de SD<sub>1</sub>; para, finalmente, tener que SD<sub>1</sub> depende de O.

Otra manera de expresar estas mismas relaciones es por medio del *syntactic bracketing* (Pustejovsky y Stubbs, 2013), que es el uso de paréntesis cuadrados para representar la estructura del enunciado en lugar de utilizar árboles, como se muestra en (5):

(5) [O [SD [D [El]] [SN [N [aro]] [SP [P [de]] [SD [D [la]] [SN [N [bicicleta]]]]]]]] [SV [V [rodaba]] [SP [P [sin]] [SN [N [problemas]]]]]]]

En cuanto a la anotación semántica, existen dos tipos característicos: la tipificación semántica y el etiquetado de roles semánticos. El primero tiene que ver con la etiqueta que se le asigna a la palabra o a la oración con el identificador que indica su significado. Estas etiquetas pueden ser en texto libre, basadas en vocabularios compartidos o basadas en identidad semántica (Hinze, Heese, Schlegel y Paschke, 2019). Las de texto libre aplican para etiquetas, frases u oraciones que se asignan a documentos, textos o imágenes, y que requieren de algoritmos lingüísticos sofisticados para su procesamiento automático. Las basadas en vocabularios compartidos se refieren al uso de vocabularios controlados, como una lista de categorías o a un esquema de clasificación existente. En cuanto a las basadas en identidad semántica, estas tienen que ver con el uso de ontologías (ver más en Pustejovsky y Stubbs, 2013). En la presente tesis no de etiquetará el significado de las palabras, sino que se limitará a trabajar específicamente con los roles temáticos.

El etiquetado de los roles temáticos, por su lado, conlleva identificar qué tipo de función semántica el verbo principal de la oración ha asignado a sus constituyentes (Pustejovsky y Stubbs, 2013). La anotación semántica representa las bases para los sistemas de pregunta-respuesta, en tanto que responde a las preguntas de qué, quién, cómo, cuándo, dónde y con qué de las oraciones. En los siguientes párrafos se profundizará sobre este tipo de anotación.

### 1.7.3 Roles temáticos

En la comunidad lingüística ha sido difícil definir una lista de roles temáticos, independientemente de la lengua que se estudie, porque no se ha llegado a un acuerdo rotundo que la justifique (Culicover y Jackendoff, 2005; Butt, 2006; Pustejovsky y Stubbs, 2013; Bresnan, Asudeh, Toivonen y Wechsler, 2016). Sin embargo, existen listas de roles temáticos que han sido recurrentes en la literatura. Algunos ejemplos de catálogos representativos se dan en la Tabla 2, todos ordenados en orden de prominencia por su autor:

**Tabla 2**

*Catálogo de roles semánticos*

	Fillmore (1968)	Jackendoff (1972)	Grimshaw (1990)	Butt (2006)	Pustejovsky y Stubbs (2013)	Bresnan et al. (2016)	Proto-roles de PropBank, Dowty (1991) y Moreda (2008)
Roles	Agente	Agente	Agente	Agente	Agente	Agente	<b>Arg0:</b> Agente
	Experimentador	Meta	Experimentador	Beneficiario	Tema / figura	Beneficiario	<b>Arg1:</b> Objeto directo/ Tema/ Paciente
	Instrumento	Fuente ( <i>source</i> )	Fuente ( <i>source</i> )	Recipiente/ experimentador	Experimentador	Experimentador/ meta	<b>Arg2:</b> Objeto indirecto/ Beneficiario/ Instrumento/ Atributo/ Estado final/ Extensión
	Paciente	Locativo	Locativo	Instrumento	Fuente ( <i>source</i> )	Instrumento	<b>Arg3:</b> Punto de partida, origen/ Beneficiario/ Instrumento/ Atributo
	Meta	Tema	Tema	Tema / paciente	Recipiente	Paciente / tema	<b>Arg4:</b> Punto de llegada, destino
	Fuente ( <i>source</i> )			Locativo	Paciente	Locativo	
	Locativo				Instrumento		
					Locativo		

Los proto-roles fueron originalmente introducidos por Dowty (citado por Pustejovsky y Batiukova, 2019) y expresan generalizaciones de los roles temáticos basados en rasgos léxicos. Estos se dividen en ‘Proto-agente’ y ‘Proto-paciente’. Los proto-agentes presentan características

de los sujetos de los verbos transitivos causativos prototípicos, como construir, matar, comer, etc. (Pustejovsky y Batiukova, 2019), y los proto-pacientes se asocian con el objeto directo de los verbos transitivos causativos prototípicos. A la hora de asignar uno de estos proto-roles a los argumentos del verbo, se siguen los siguientes pasos (Pustejovsky y Batiukova, 2019, pág. 223):

- a) El argumento con el mayor número de propiedades de proto-agente será el sujeto del enunciado.
- b) El argumento con el mayor número de propiedades de proto-paciente será el objeto directo.
- c) Dos argumentos con un número similar de propiedades de proto-agente, entonces el sujeto o el objeto directo pueden ser el agente, como en casos de los pares de verbos gustar/complacer y temer/asustar, que contienen verbos con significados similares, pero con estructuras-a diferentes (ver más sobre estos casos en Pustejovsky y Batiukova, 2019).

Pustejovsky y Batiukova (2019) opinan que los roles temáticos no juegan un papel importante en la teoría lingüística generativa, porque no se han considerado como una noción básica, sino como un derivado. Por esta razón, Butt (2006) expresa que los proto-roles han venido a ser una solución para mejorar la insatisfacción general que la comunidad lingüística ha tenido en relación con la teoría estándar, en cuanto al análisis de los roles temáticos. De aquí que se haya optado por aplicar los proto-roles para simplificar los análisis semánticos y las anotaciones automáticas, puesto que permiten analizar qué tan propensos son los argumentos verbales para encajar en un rol específico (Moreda, 2008).

No obstante, la investigadora de esta tesis considera que dividir los argumentos verbales entre “proto-agentes” y “proto-pacientes” no aporta conocimiento lingüístico de gran valor para el entrenamiento de ML, porque despoja a los verbos de su variedad y riqueza semántica en pro de simplificar la clasificación y la automatización.

Además de los proto-roles en la Tabla 2, el *English Propbank Annotation Guidelines* (Bonial, Hwang, Bonn, Conger, Babko-Malaya y Palmer, 2015) ofrece una lista de etiquetas para anotar consistentemente tanto los argumentos como los modificadores verbales. Las etiquetas que proponen Bonial et al. (2015) para los modificadores incluyen: comitativo (COM) para compañía; locativo (LOC) para lugar; direccional (DIR) para origen; meta (GOL); manera (MNR); temporal (TEMP) para tiempo; extensión (EXT) para alcance; recíproco (REC) para los pronombres

recíprocos y los reflexivos; finalidad (PRP); causativo (CAU) para las forma causativas; discurso (DIS) para los conectores; modal (MOD) para las formas auxiliares del verbo; adverbiales (ADV) para ciertos complementos circunstanciales; y negación (NEG), entre otras etiquetas pertinentes al idioma inglés. Las etiquetas que la investigadora menciona aquí tienen más relevancia para el español y para el tipo de anotación que se propuso en los objetivos.

De la Tabla 2, se desprende que, dependiendo del tamaño del corpus, del tiempo y de los recursos que disponga el investigador así también determinará qué tan detallada y variada sea la anotación de los roles semánticos que se lleve a cabo (Hinze et al., 2019). A continuación, se explicarán las relaciones sintáctico-semánticas de los argumentos verbales, desde la perspectiva de la Gramática Léxico-Funcional (LFG), y cómo esta gramática asigna los roles temáticos para los argumentos del predicado (PRED).

#### ***1.7.4 Relaciones sintáctico-semánticas de los argumentos verbales desde la Gramática Léxico-Funcional***

La Gramática Léxico-Funcional (LFG) tuvo sus inicios en los años 70, cuando Joan Bresnan comenzó a cuestionar la viabilidad de la Gramática Transformacional y tomó en consideración las ideas de Kaplan sobre la psicolingüística y el modelado computacional (Butt, 2006). Las ideas que dieron origen al LFG mantienen todavía su validez, porque su acercamiento lógico-matemático permite que los sistemas computacionales comprendan las estructuras de los análisis lingüísticos por medio de ecuaciones funcionales fácilmente interpretables (Butt, 2006; Bresnan et al., 2016; Pustejovsky y Batiukova, 2019). A continuación, se dará una breve reseña de qué elementos caracterizan al LFG y por qué se consideraron estos como los apropiados para desarrollar la presente investigación.

Como gramática lexicalista, el LFG defiende el papel crucial que juega el léxico para construir las estructuras de los dos niveles principales, el interno y el externo, de frases pertenecientes a una lengua X (Butt, 2006; Bresnan et al., 2016; Pustejovsky y Batiukova, 2019). Los niveles de estructuras que el LFG ha trabajado son la estructura-a (estructura argumental, relacionada con la teoría- $\theta$  y ligamento) que da lugar al análisis de los roles semánticos; la estructura-f para la representación de las funciones gramaticales (también conocida como estructura interna); la estructura-c para la representación de los constituyentes sintácticos (o estructura externa); la estructura-m para la representación morfológica, y la estructura-p para la representación fonética. De estas cinco estructuras, las 3 primeras son las más desarrolladas en la

literatura (Bresnan et al., 2016; Pustejovsky y Batiukova, 2019) y, para la presente investigación, interesan, por supuesto, las estructuras-a, -c y -f.

Los análisis lingüísticos que se construyen sobre LFG tienen la facilidad de estar estructuralmente organizados para implementarse en ML y en LComp, porque sus bases se formaron con esta idea. Para la presente tesis, es importante contar con un soporte teórico que permita etiquetar las Funciones Gramaticales y las Funciones Temáticas con precisión y con un formato fácilmente comprensible por sistemas computacionales, para que luego, eventualmente, se puedan llevar a cabo mapeos que rastreen la información lingüística por medio de algoritmos computacionales. Con este trabajo la investigadora no busca refutar, ni criticar perspectivas teóricas, mucho menos crear nuevas visiones, sino que tiene en mente aportar al análisis de los roles semánticos en español, un área que carece de una exploración detenida. A continuación, se explican las estructuras de los niveles pertinentes para entender las relaciones sintáctico-semánticas entre los verbos y los participantes de la oración, según LFG.

#### **1.7.4.1 La estructura-c**

La estructura-c es la representación arbórea del orden de las palabras y de la organización jerárquica de los constituyentes de una oración, porque únicamente cifra (*encode*) las dos restricciones configuracionales: la linealidad y la dominancia (Pustejovsky y Batiukova, 2019). Para crear las estructuras-c, existen dos convenciones de notación que se deben seguir para mantener su regularidad, y estas son la forma de las reglas sintácticas y la forma de las unidades léxicas (Wescoat, 1987; Bresnan et al., 2016).

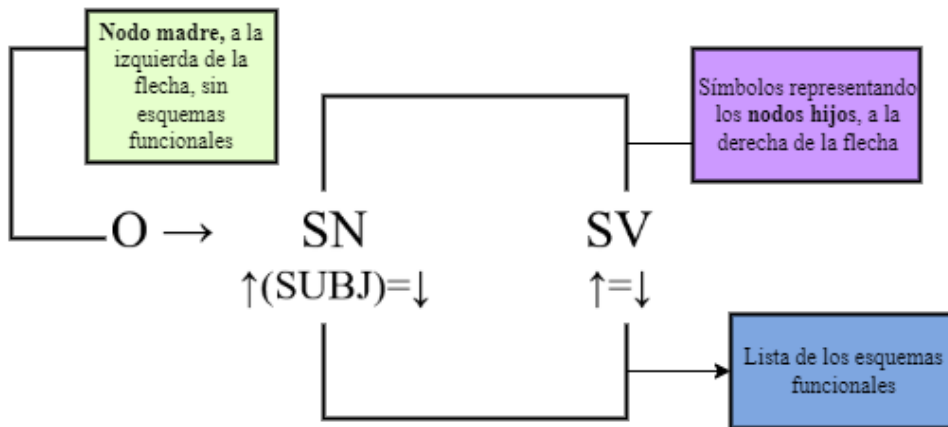
LFG toma como base una versión de la Teoría de la X' y las reglas de reestructura libres de contexto del tipo  $O \rightarrow SN\ SV$ , enriquecidas con expresiones regulares matemática y computacionalmente rastreables (Wescoat, 1987; Butt, 2006; Bresnan et al., 2016). Este tipo de reglas, de herencia chomskiana, son un estándar en la LComp (Wescoat, 1987; Bresnan et al., 2016). Wescoat (1987) llama a la estructura-c «estructura anotada de los constituyentes» (pág. 1), en donde el término ‘anotada’ tiene que ver con que los nodos del árbol sintáctico contienen anotaciones con expresiones regulares, conocidas como “*functional schemata*” o esquemas funcionales, que luego servirán para construir la estructura-f o estructura interna.

Las convenciones de notación tienen que ver con la forma de las reglas sintácticas de reescritura y con la forma de las unidades léxicas. En LFG, las reglas constan de un nodo madre

(*mother node*) que se encuentra a la izquierda de la flecha y se representa sin esquemas funcionales. A la derecha de la flecha se encuentran los nodos hijos (*daughter nodes*) junto con los esquemas funcionales que le corresponden a cada nodo (Wescoat, 1987; Bresnan et al., 2016). En la Figura 9 se representan estos niveles de jerarquía o de dependencia, mismo principio que se explicó para las etiquetas POS.

### Figura 9

*Formato de las reglas en LFG*



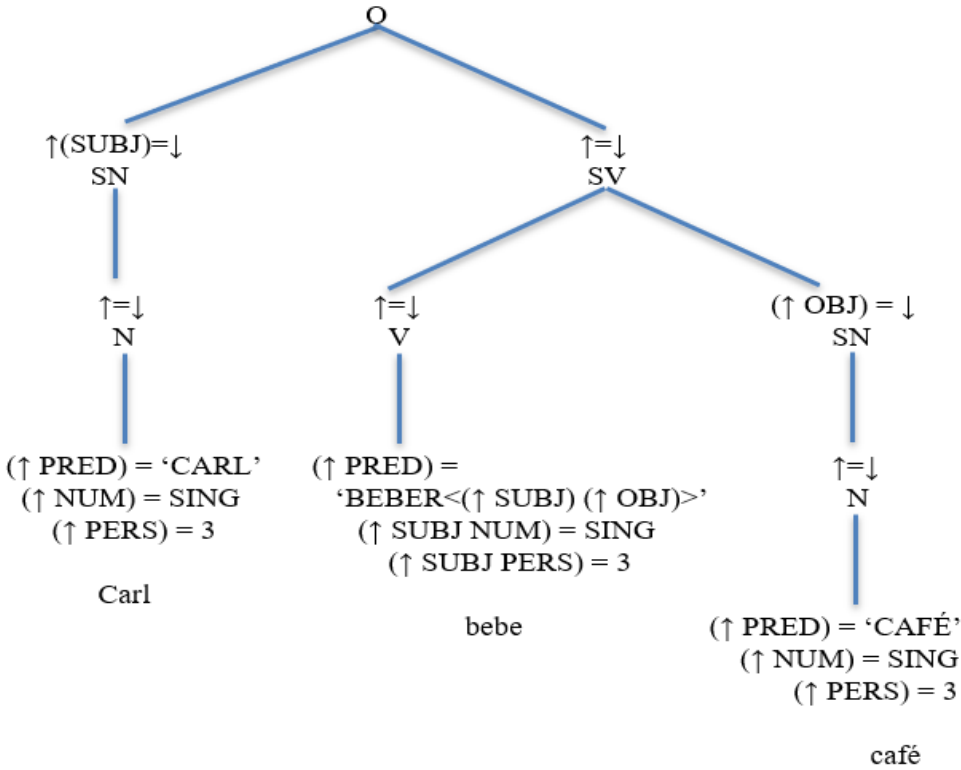
La forma de las unidades léxicas requieren de un mínimo de tres cosas: de la representación de la forma de la unidad, de su categoría sintáctica y de una lista con los esquemas funcionales. Para la frase, *Carl bebe café*, la forma de sus unidades léxicas se representaría de la siguiente manera en (6):

- (6) Carl bebe café.
- |      |   |                                     |
|------|---|-------------------------------------|
| Carl | N | (↑ PRED) = 'CARL'                   |
|      |   | (↑ NUM) = SING                      |
|      |   | (↑ PERS) = 3                        |
| bebe | V | (↑ PRED) = 'BEBER<(↑ SUBJ)(↑ OBJ)>' |
|      |   | (↑ SUBJ NUM) = SING                 |
|      |   | (↑ SUBJ PERS) = 3                   |
| café | N | (↑ PRED) = 'CAFÉ'                   |
|      |   | (↑ NUM) = SING                      |
|      |   | (↑ PERS) = 3                        |

Una vez establecidas las formas de las unidades léxicas en (6) como anotaciones, estas deben integrarse al árbol sintáctico que representa la estructura externa o estructura-c, como se aprecia en la Figura 10:

**Figura 10**

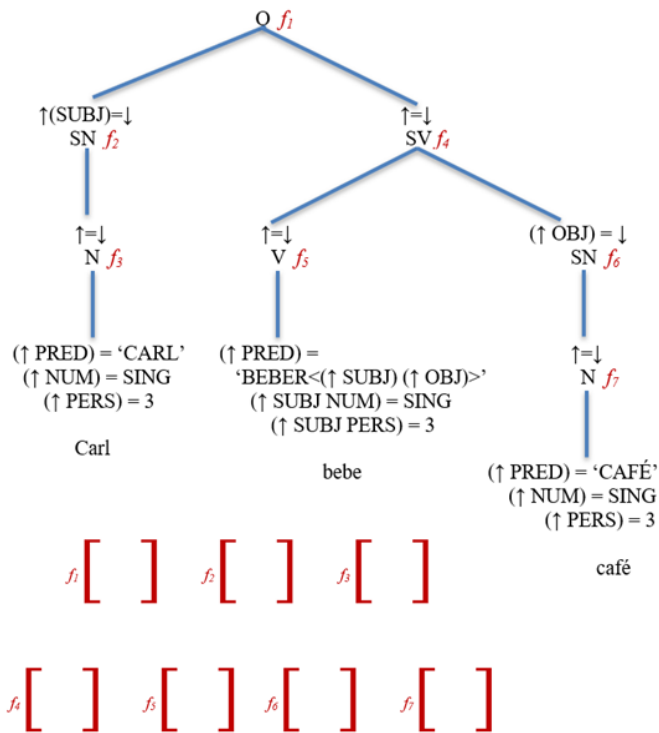
*Árbol de constituyentes anotado con esquemas funcionales*



Cuando se completa la adición de los esquemas funcionales al árbol sintáctico, se procede a lo que se conoce como ‘instanciación’: transformar los esquemas funcionales en ecuaciones funcionales, las cuales son expresiones escritas en un lenguaje formal con especificaciones completas que serán la base para construir las estructuras-f posteriormente (Wescoat, 1987; Bresnan et al., 2016). De manera que, en la figura 11, se encuentran siete estructuras-f asociadas con los nodos de la frase en (6):

## Figura 11

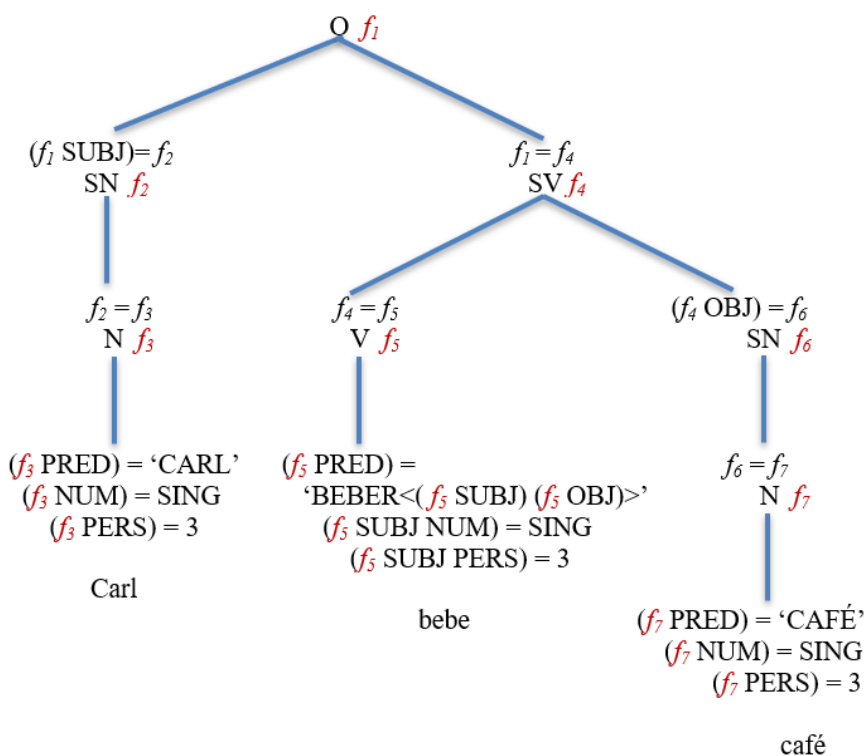
Asociación entre los nombres de las estructuras-c y sus nodos



En la instanciación se procede a asignar ecuaciones funcionales a cada metavariable que se encuentra en el árbol sintáctico. Los nombres de cada variable son arbitrarios, pero deben ser únicos; no obstante, en LFG ha sido tradicional emplear la letra  $f$  seguida de un número,  $f_{123}$ , de manera que se asocia una estructura- $f$  con un nodo por medio de una variable  $f_n$ . En la Figura 12 se muestran las ecuaciones funcionales para cada nodo.

**Figura 12**

*Consolidación del algoritmo de instanciación*



Siguiendo la explicación de Pustejovsky y Batiukova (2019), las ecuaciones funcionales son restricciones formalizadas que controlan la forma en que se comunica la información dentro de la estructura-c, y entre la sintaxis y la estructura-f; y su forma es del tipo  $(f_n \text{ ATT}) = \text{VAL}$ , en donde ATT se refiere a ‘atributo’ y VAL se refiere al ‘valor’, por lo que una estructura-f se compone de pares de atributos-valores ordenados en matrices.

Ahora bien, los verbos codifican información sobre la estructura de sus argumentos, como la función gramatical de cada argumento, los valores de concordancia, tiempo, aspecto, las propiedades de asignación de caso, etc., de manera que la entrada léxica del verbo en el árbol pasará la información ‘atributo-valor’ hacia arriba, hacia el nodo incrustado arriba, en donde, si es núcleo del sintagma, pasará recursivamente la información hacia la parte alta del árbol o ‘nodo raíz’, ‘O’. Es así como O cargará con el valor de la estructura-f resultante de V para toda la oración. Se volverá a esto en unos párrafos.

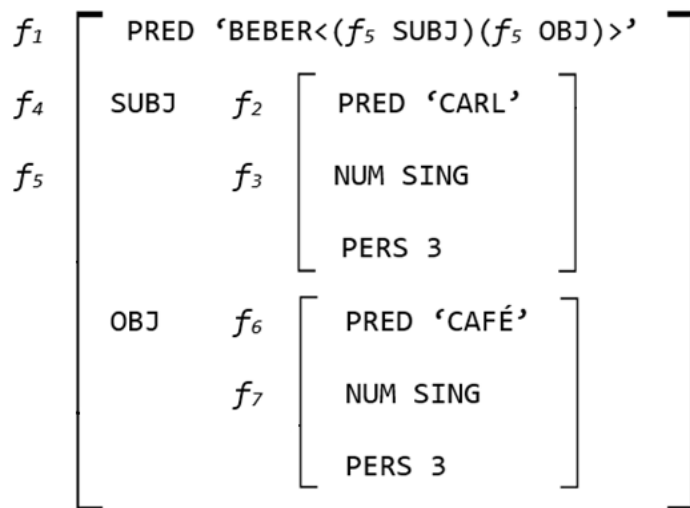
### 1.7.4.2 La estructura-f

Para construir la estructura-f se toman en cuenta únicamente las descripciones funcionales que se encuentran en las ecuaciones  $f_n$  de la Figura 13 y el léxico creado para cada oración, por lo que la información categorial no se transmitirá a esta estructura (Wescoat, 1987; Butt, 2006; Bresnan et al., 2016). En este punto se debe recordar que en LFG no hay una derivación de una estructura-c a una estructura-f o viceversa, sino que estos dos niveles de representación se limitan mutuamente y su relación se expresa en términos de una proyección matemática, en donde se puede ‘ver’ la estructura-f desde la estructura-c y la estructura-f refiere en retorno a la estructura-c (Butt, 2006). Por esta característica particular, en LFG no hay un nivel primario ni se pierde información, porque LFG no adopta derivaciones de una estructura a la otra, como ocurre en otras corrientes teóricas (Butt, 2006).

La estructura-f se basa en dos características principales: su forma gráfica y su condición de singularidad (*uniqueness*) (Wescoat, 1987). La forma de la estructura-f tiene dos columnas de entradas encerradas en corchetes o paréntesis cuadrados [], en donde la columna de la izquierda contiene los atributos, mientras que la de la derecha tiene los valores de dichos atributos. El orden de la lista de las descripciones funcionales no tiene mayor significancia, siempre y cuando los pares de atributos-valores estén organizados horizontalmente en la tabla de la matriz, como se perfila en la Figura 13:

**Figura 13**

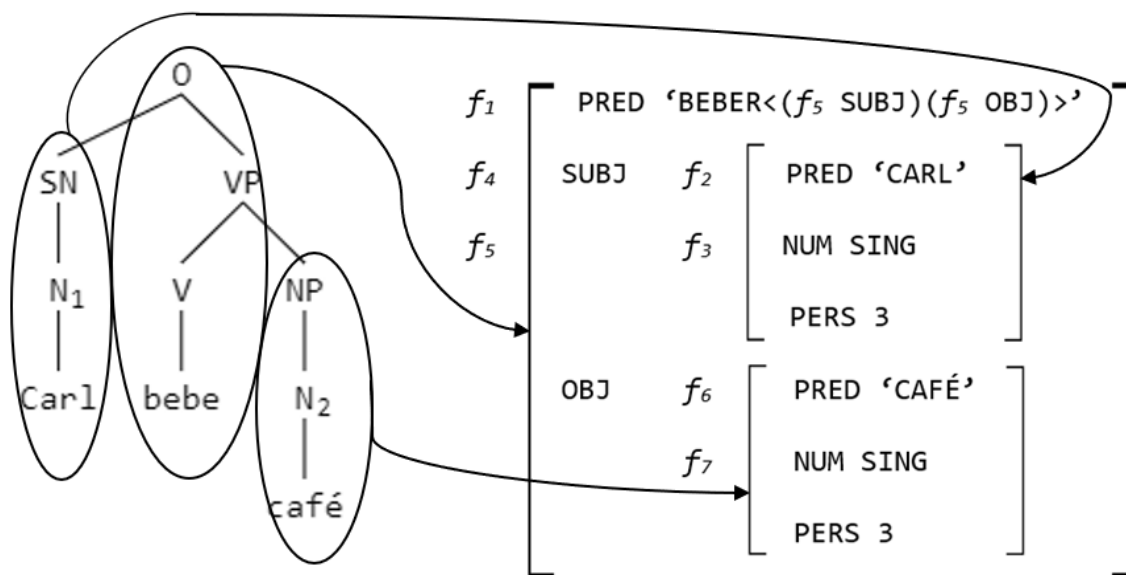
*Representación de la estructura-f*



En la Figura 14, a la izquierda, afuera de los paréntesis cuadrados, se encuentra el nombre de cada estructura-f ( $f_1$ ,  $f_2$ ,  $f_3$ , etc.), y, en cuanto a la singularidad, se exige que cada atributo tenga un solo valor por estructura-f (Wescoat, 1987; Butt, 2006; Bresnan et al., 2016), a saber: PRED es ‘CARL’ para  $f_2$ , mientras que PRED es ‘CAFÉ’ para  $f_6$ . Es preciso recordar que se ha venido trabajando con una versión simplificada de las reglas de reescritura y que por esta razón no se han integrado todos los posibles atributos que se precisan en un análisis lingüístico más completo, como aspecto, modo o tiempo verbal. Resumiendo, en la Figura 14 se representa la correspondencia entre la estructura-c y la estructura-f:

**Figura 14**

*Correspondencia entre la estructura-c y la estructura-f*



De acuerdo con Butt (2006), la estructura-f contiene las relaciones gramaticales de la oración analizada. Estas relaciones gramaticales son las relaciones entre los argumentos y los elementos predicacionales y son representadas por medio de funciones gramaticales (FG) que indican el estado funcional de cada argumento. Una de las principales motivaciones de LFG es construir una teoría de funciones gramaticales, y es por medio de la estructura-f que las relaciones entre estas funciones se pueden mostrar (Wescoat, 1987). LFG adopta una organización jerárquica de las FG y en (7) se presenta dicha jerarquía:

- (7) SUBJ OBJ OBJ<sub>θ</sub> OBL(ícuo)<sub>θ</sub> COMP(lemento) XCOMP(lemento) ADJUNTO

Butt (2006) explica que, de acuerdo con (7), COMP y XCOMP, respectivamente, representan los argumentos de cláusulas finitas (en donde todos los argumentos se realizan internamente a la cláusula) y las no finitas (con cláusulas incrustadas abiertas no finitas).  $OBJ_{\theta}$  se reserva para el objeto indirecto o complemento dativo, el OBL para las frases preposicionales, también argumentos verbales. Los subíndices  $\theta$  en OBJ y OBL indican que este tipo de FG son sensibles a la información del rol temático, porque estas FG tienden a reflejar tipos específicos de roles, como meta o beneficiario para el  $OBJ_{\theta}$ , mientras que SUBJ y OBJ no son tan selectivos en cuanto a cuál rol temático se les asigna (Butt, 2006; Bresnan et al., 2016). Esta estructura es de gran importancia para la presente investigación, porque permite definir las FG adecuadamente, para relacionarlas luego con las relaciones temáticas entre los participantes de la oración y el verbo principal en la estructura-a o estructura de argumentos. En el apartado sobre estructura argumental se profundizará más al respecto.

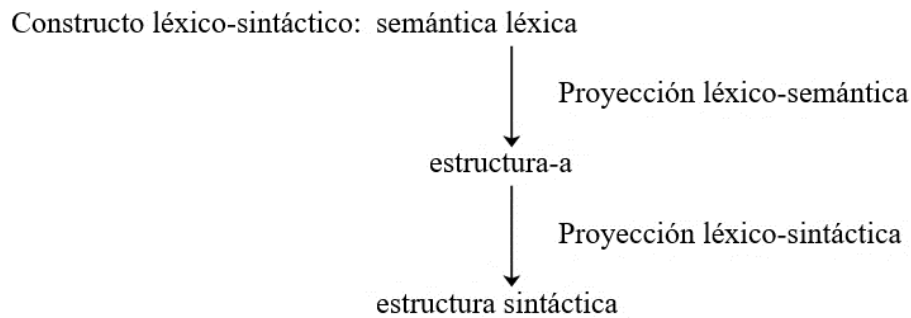
#### 1.7.4.3 La estructura-a

La estructura-a o estructura argumental es «una interfaz entre la semántica y la sintaxis del predicado [...], (porque) [...] codifica información léxica sobre el número de argumentos, su clase sintáctica, y la organización jerárquica necesaria para el mapeo a la estructura sintáctica» (Bresnan et al., 2016, pág. 326). Generalmente, se toma el caso para los verbos, aunque la estructura argumental puede verse como una propiedad de todas las categorías mayores, como explican Pustejovsky y Batiukova (2019). No obstante, para efectos de esta tesis, se tomarán las relaciones argumentales de los eventos verbales únicamente.

Bresnan et al. (2016) sostienen que la estructura-a es un constructo léxico-sintáctico, porque las estructuras argumentales «deben tener suficiente información sintáctica para soportar el mapeo directo a funciones superficiales» (pág. 329). Pustejovsky y Batiukova (2019) definen la estructura-a como la especificación de tanto el número como el tipo de argumentos lógicos para un predicado y que su noción proviene del álgebra, en cuanto a los conceptos de relación y funciones matemáticas. En la Figura 15 se muestra este constructo, en donde la proyección léxico-semántica se basa en el *lexical mapping theory* (LMT) para proyectar estructuras-f reducidas a partir de estructuras-a, y este mapeo se da entre los roles temáticos de mayor prominencia en la estructura-a y las FG más prominentes en la estructura-f, por medio de principios generales (Bresnan et al., 2016) que se explicarán en los siguientes párrafos.

## Figura 15

### Constructo léxico-sintáctico



Ya se explicaron los conceptos de estructura-c y estructura-f en los párrafos anteriores. En LFG, la estructura-a consiste en un predicado con los roles de sus argumentos, el orden que representa la prominencia relativa de los roles y la clasificación sintáctica de cada rol por medio de rasgos o características de restricción. Ejemplos de estructuras-a se dan en (8) y (9), en donde para (8) ‘poner’ es el predicado y los roles de sus argumentos son *agente*, *tema* y *locativo*, en orden de prominencia, y [-o], [-r] y [-o] son los rasgos de restricción para cada uno de sus roles:

(8) *poner* < *agente tema locativo* >

[-o] [-r] [-o]

(9) *golpear* < *agente paciente* >

[-o] [-r]

Con (7) se explicó que LFG emplea una organización jerárquica para las funciones gramaticales y que estas se representan únicamente en el nivel interno o estructura-f, por ser parte de la descripción lingüística. La organización jerárquica de las FG comparte una relación estrecha con los roles temáticos, porque los roles temáticos se organizan también jerárquicamente en LFG (Butt, 2006; Bresnan et al., 2016), para que el LMT explicado arriba se pueda aplicar. Por ejemplo, en (8) el rol temático *agente* es el más prominente y este corresponderá con la función gramatical de SUBJ, al igual que en (9). En (10) se muestra la jerarquía de prominencia para los roles temáticos que emplean Bresnan et al. (2016):

(10) agente > beneficiario > experimentador/meta > instrumento > paciente/tema > locativo



+ prominente

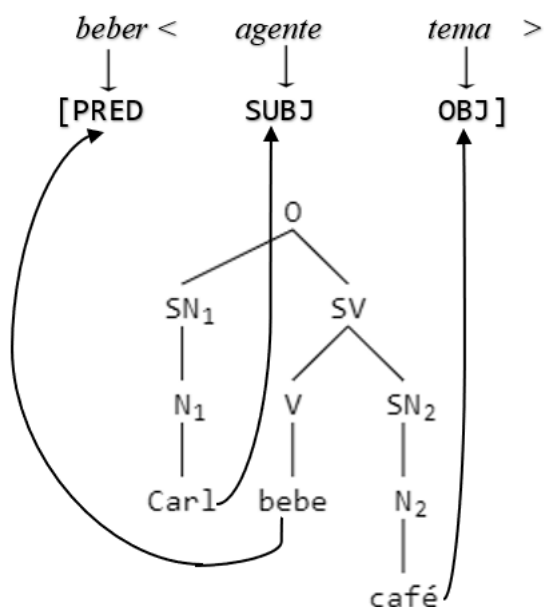
- prominente

Para efectos prácticos, en la Figura 16 se ilustra la correspondencia entre la estructura-c, la estructura-f y la estructura-a para el ejemplo de la oración *Carl bebe café*, tomando como base el

estructura dado en la Figura 15. Aquí, el SUBJ, la FG más prominente para esta frase concuerda con el rol temático más prominente *agente*; mientras que la FG menos prominente OBJ concuerda con el rol menos prominente *tema* para este ejemplo:

**Figura 16**

*Correspondencia entre las estructuras -c, -f y -a para Carl bebe café*



#### 1.7.4.4 El número de valencia

En este punto es importante recordar que los verbos describen eventos, y la naturaleza de estos eventos determina cuáles participantes están cifrados como argumentos verbales, cuál es la interacción entre ellos y qué roles temáticos juegan en el evento. Los verbos poseen número de valencia, el cual indica el número de argumentos y el tipo semántico de argumentos que se relacionan con el evento (Pustejovsky y Batiukova, 2019). Para garantizar la asociación correcta entre los argumentos y el predicado, en LFG se aplican dos principios del modelo de *Government and Binding* (GB): el *Uniformity of Theta Assignment Hypothesis* (UTAH), que restringe el mapeo entre las alternaciones entre formas transitivas e intransitivas (voz activa/pasiva), y el Criterio  $\theta$  (theta).

El criterio theta requiere que a cada argumento de la estructura temática proyectada por el verbo se le asigne un rol- $\theta$  y que a cada rol- $\theta$  se le asigne un solo argumento (Pustejovsky y Batiukova, 2019). Para el caso del sujeto, el rol- $\theta$  se le asigna de una manera distinta a la de los

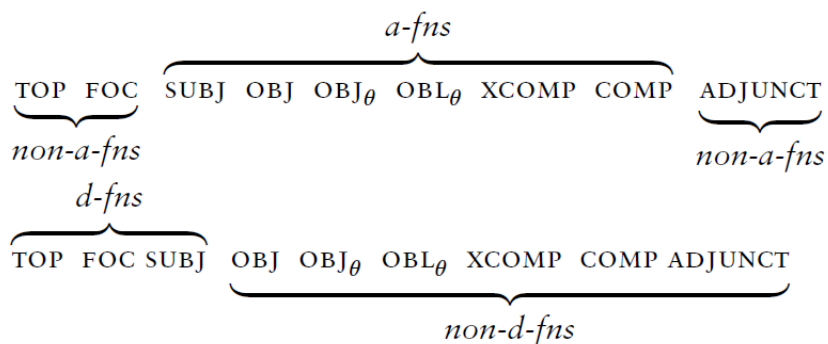
otros argumentos, puesto que el sujeto es un argumento externo con respecto del SV, y se asume que su rol- $\theta$  se lo asigna el SV como un todo y que no proviene directamente del verbo en sí, mientras que los otros argumentos son internos y obtienen su rol- $\theta$  directamente del verbo (Pustejovsky y Batiukova, 2019).

En cuanto a argumento, este se puede definir, en un sentido amplio, como todos los componentes de un predicado, entiéndanse argumentos y adjuntos, excluyendo al núcleo, y, en un sentido más estricto, como a los componentes requeridos únicamente, dejando de lado a los adjuntos (Pustejovsky y Batiukova, 2019). Para efectos de esta tesis, se trabajará con el sentido amplio de esta definición, porque interesa tomar en cuenta tanto los argumentos requeridos estrictamente, como ciertos adjuntos necesarios para responder interrogantes como cuándo y dónde.

Bresnan et al. (2016) distinguen entre funciones de argumento (*a-fns*) y de no argumento (*non-a-fns*), y entre funciones discursivas gramaticalizadas (*d-fns*) y no discursivas (*non-d-fns*), como se expone en la Figura 17, en donde se puede percibir que los adjuntos (ADJUNT), el tópico (TOP) y el foco (FOC) quedan externos a la función argumental, según la definición estricta de (Pustejovsky y Batiukova, 2019). SUBJ tiene la particularidad de ser tanto una función argumental como una función discursiva gramaticalizada, y las funciones de complemento (OBJ, OBJ $\theta$ , OBL $\theta$ , XCOMP, COMP, ADJUNT) son las funciones del argumento no discursivo (Figura 17) (Bresnan et al., 2016, pág. 330).

**Figura 17**

*Distintos tipos de funciones*



Párrafos arriba se explicó sobre el número de valencia y, con este, los verbos se clasifican en avalentes o impersonales, en monovalentes o intransitivos (que a su vez se subclasifican en inergativos, con sujetos que manifiestan propiedades de sujetos de verbos transitivos; e

inacusativos, con sujetos que tienen propiedades de los objetos directos de los verbos transitivos), en bivalentes o transitivos (verbos de cambio de estado/verbos estativos), y en ditransitivos (Pustejovsky y Batiukova, 2019). A continuación, se presentan ejemplos de esta clasificación:

- Avalente: Nieva. Llueve. Truena.
- Monovalente o intransitivo:
  - Inergativo: [Isaac] corrió.
  - Inacusativo: [Putin] murió.
- Transitivo:
  - Cambio de estado: [Carl] quebró [la taza rosa].
  - Estativo: [Carl] sabía [la respuesta].
- Ditransitivo: [Carl] le dio [galletas] a [Isaac].

Bresnan et al. (2016) clasifican las funciones argumentales en clases naturales, de acuerdo con los rasgos  $[\pm o]$   $[\pm r]$ , y esta clasificación se ilustra en la Tabla 3 (Bresnan et al., 2016, pág. 330). Estos rasgos se utilizan en la estructura-a para restringir la forma en que los roles y las funciones de argumento en las estructuras-f se identifican.

**Tabla 3**

*Descomposición de características de funciones de argumento*

	$-r$	$+r$
$-o$	SUBJ	OBL <sub><math>\theta</math></sub>
$+o$	OBJ	OBJ <sub><math>\theta</math></sub>

En donde,  $[\pm r]$ : (no)*restringido*

$[\pm o]$ : (no)*objetivo*

El rasgo  $[-r]$  indica que se trata de una función irrestricta en relación con su rol semántico, porque no requiere un rol semántico. Butt (2006) y Bresnan et al. (2016) concuerdan, por un lado, con que las únicas funciones que poseen este rasgo son SUBJ y OBJ (ver la Tabla 3), mientras que OBL <sub>$\theta$</sub>  y OBJ <sub>$\theta$</sub>  son  $[+r]$ , porque sí necesitan un rol semántico. Por otro lado, el rasgo  $[-o]$  hace referencia a la función sintáctica no objetiva, la cual complementa los predicados intransitivos como N o A, y solo los SUBJ y los OBL <sub>$\theta$</sub>  son  $[-o]$ , mientras que los objetos OBJ y OBJ <sub>$\theta$</sub>  son  $[+o]$ . Esta clasificación se basa en que LMT asume que todas las lenguas naturales tienen sujeto, aunque

algunas lenguas no tengan  $OBJ_{\theta}$ . Cabe mencionar aquí que, en esta tesis, no se entrará en detalles sobre estas discusiones teóricas y se asumirá que las oraciones analizadas tendrán un sujeto.

Volviendo a la Tabla 3, los rasgos con [-] de  $o$  y  $r$ , se consideran ‘no marcados’, de lo que se obtiene que SUBJ es la FG menos marcada, y  $OBJ_{\theta}$ , es la más marcada. Así las cosas, el orden de las FG, de menos marcada a más marcada, sería la presentada en (12):

$$(11) \quad \text{SUBJ} > \text{OBJ}, \text{OBL}_{\theta} > \text{OBJ}_{\theta}$$

Tomando de nuevo la Tabla 3 en consideración, si un rol temático en la estructura-a tiene el rasgo [- $o$ ], este no puede corresponder con un objeto en la estructura-f, mientras que un rol con el rasgo [- $r$ ] puede corresponder con un sujeto o con un objeto. Butt (2006) y Bresnan et al. (2016) indican que los principios básicos para determinar cuáles rasgos sintácticos irán en la estructura-a se basan en la semántica léxica subyacente de los roles:

(12) Clasificación semántica por función de los roles de la estructura-a:

- Roles tipo paciente:  $\theta$   
[- $r$ ]
- Roles secundarios tipo paciente:  $\theta$   
[+ $o$ ]
- Otros roles semánticos:  $\theta$   
[- $o$ ]

Bresnan et al. (2016) afirman que con la clasificación en (13) se pueden distinguir diferentes tipos de ‘sujetos lógicos’ ([- $r$ ] o [+ $o$ ]), de los no-objeto ([- $o$ ]) y, junto con los principios de relación (*Mapping Principles*) y la condición de buena formación (*Wellformedness Conditions*) se regula la posible correspondencia entre los roles y las funciones gramaticales, como se aprecia en (13):

(13)	agente	tema/paciente	locativo
	[- $o$ ]	[- $r$ ]	[- $o$ ]
	SUBJ/OBL	SUBJ/OBJ	SUBJ/OBL

Siguiendo a Butt (2006) y a Bresnan et al. (2016), los *Mapping Principles* (MP) y la *Wellformedness Condition* (WC) se enumeran en (15) y (16), en donde  $\theta$  indica el rol temático y  $\hat{\theta}$  representa el argumento más alto en la jerarquía de los roles temáticos:

(14) *Mapping Principles*

a. Roles de sujetos:

- i. A  $\hat{\theta}$  se le asigna la función de SUBJ; si no:

[-o]

- ii. A  $\theta$  se le asigna la función de SUBJ

[-r]

A los otros roles se le asignará a la función más baja compatible en la jerarquía de marcas  $[\pm]$ , en donde se sabe que el sujeto es el menos marcado:  $\vartheta$

$$\text{SUBJ} < \text{OBJ}, \text{OBL}_{\theta} < \text{OBJ}_{\theta}$$

(15) *Wellformedness Condition:*

- a. Condición de sujeto: toda forma léxica verbal debe tener un sujeto.  
b. Biunicidad función-argumento: cada estructura-a debe estar asociada con una única función gramatical y viceversa.

Con MP y WC, LFG se asegura de que cada rol temático asignado sea válido en cuanto a la relación que mantiene con la función gramatical de dicho argumento y conforme a las jerarquías de ambas estructuras: la de las FG y la de los roles temáticos. Estos conceptos producirán etiquetados fidedignos en la presente investigación, los cuales se podrán emplear y aplicar en entrenamiento automático para ML.

En este apartado se han desarrollado las bases teóricas pertinentes para llevar a cabo los análisis sintáctico y semántico, imprescindibles para anotar los roles temáticos en español. Se explicaron los tipos de anotación lingüística necesarios para los corpus lingüísticos, como el POS y los varios catálogos de roles temáticos pertinentes, de los cuales la investigadora de esta tesis prefiere los empleados por Bresnan et al. (2016), porque en LFG no se recurre a listas de roles temáticos interminables, sino que clasifica los argumentos dependiendo de su función dentro del predicado con una lista de roles consistentes, de manera que se pueden llevar a cabo etiquetados acertados, con jerarquías lógicas para tomar decisiones congruentes. LFG tampoco se limita a emplear proto-roles que terminan generando vaguedades en cuanto a qué roles realmente desempeñan los argumentos verbales en los corpus analizados y que producirán entrenamientos automáticos de carácter insustancial.

Este apartado explicó, a grandes rasgos, cómo LFG trabaja las estructuras-c, -f y -a para interpretar las relaciones sintáctico-semánticas entre las funciones gramaticales de los argumentos y los roles temáticos que se le asignan a estos. La investigadora de esta tesis estima de gran valor las estructuras-c, -f y -a de LFG para llevar a cabo la aproximación lingüística en el diseño de un corpus anotado en español sobre COVID-19 para sistemas de pregunta-respuesta. Las razones son las siguientes.

En primer lugar, porque su mecanismo de carácter accesible y su perspectiva lógico-matemática simplifica la anotación lingüística de corpus escritos, en conjunto con LC como metodología. En segundo lugar, porque al emplear LFG para construir los análisis sintácticos y semánticos para esta tesis se producirán etiquetas con respaldo teórico que garantizan un material válido para entrenar algoritmos de aprendizaje automático posteriormente. Y, en tercer lugar, porque la decisión entre un rol temático u otro para un determinado argumento se resolverá aplicando los métodos de rasgos  $[\pm o]$   $[\pm r]$  y no subjetivamente por la anotadora, como se explicó en párrafos anteriores.

En adición, se revisaron los términos del valor de valencia para los eventos verbales, con los que se clasifican los verbos dependiendo de la cantidad y tipo de argumento que estos requieran, aportando al etiquetado mayor grado de precisión. A pesar de que Bresnan et al. (2016) distinguen entre funciones de argumento (*a-fns*) y de no argumento o adjuntos (*non-a-fns*), la investigadora de esta tesis considera preciso no hacer esta distinción en el análisis textual porque se requieren los adjuntos de tiempo para responder las interrogantes de cuándo. En LFG sí se etiqueta el rol ‘locativo’ para lugar, pero excluye ‘tiempo’ de su listado. Por esta razón se analizarán tanto los argumentos verbales como los adjuntos que aparezcan en las oraciones del corpus. En el siguiente apartado se expondrá la metodología que seguirá la investigadora.

## **1.8 Metodología**

En este último apartado del Capítulo I se detallará la metodología que se siguió para cumplir los objetivos de la investigación. En los siguientes párrafos se presentarán, el problema lingüístico que apunta esta investigación en 1.8.1, las características de esta investigación en 1.8.2, la descripción del corpus en 1.8.3 y el procesamiento del *Corpus COVID-19* en 1.8.4, con todas las tareas que se requirieron para la anotación efectiva de los roles temáticos. Además, se explicará

la estadística empleada para evaluar el acuerdo entre los anotadores en 1.8.4.4. También se describirán las herramientas elegidas para conseguir cada uno de los objetivos planteados.

### ***1.8.1 El problema lingüístico***

NLP y *Machine Learning* requieren corpus que hayan sido anotados con la información lingüística pertinente para alcanzar la meta buscada en el entrenamiento. En el caso de los sistemas de pregunta-respuesta, esta necesidad es todavía más imperante. El análisis de roles temáticos en español ha sido un tema poco tratado en la literatura y es por esta razón que la presente tesis tiene como objetivo ofrecer una propuesta de anotación de roles temáticos en español sobre un corpus con un tema actual, el COVID-19, para que pueda ser empleado, eventualmente, por investigadores del área de Lingüística Computacional o de Informática para entrenar *chatbots* que respondan preguntas sobre esta pandemia o sobre cualquier otra eventualidad.

Es importante recordar que este problema lingüístico también tiene sus bases en que no ha existido un acuerdo entre lingüistas en cuanto a una lista única de roles temáticos, como se mencionó en el marco teórico. Será difícil concordar en una lista mínima o máxima, porque siempre dependerá de qué tan exhaustivo sea el estudio que se lleve a cabo. La investigadora de esta tesis no pretende definir una lista de roles a seguir, sino que propone un modelo de etiquetado que sea replicable en otros corpus y con otros anotadores, sean estos humanos o automatizados.

### ***1.8.2 Tipo de investigación***

La perspectiva lingüística desde la que se aborda esta investigación es la gramática generativa, específicamente, la Gramática Léxico-Funcional. Sin embargo, la Lingüística de Corpus (LC) es la metodología que guía este trabajo. Guillermo Rojo (2021) señala que la definición de LC ha sido debatida en la literatura entre si se trata de una nueva teoría, una nueva (sub)disciplina de la lingüística o una metodología. Rojo explica que no es una nueva teoría, porque diversos marcos teóricos pueden ser empleados en las investigaciones que utilicen la LC, y que tampoco es una nueva (sub)disciplina de la lingüística, porque la LC se puede aplicar en estudios gramaticales, históricos, lexicográficos o semánticos, por ejemplo; y, en cuanto a si es una metodología, Guillermo Rojo acepta que sí lo es, aunque él cree que es mucho más que eso cuando postula que la LC se trata de:

una aproximación al estudio de los hechos lingüísticos de orientación empírica y basada en el análisis detallado de gran cantidad de datos (los corpus), con lo que queda patente su oposición tanto a la lingüística racionalista como a la descriptiva tradicional. (Rojo, 2021, pág. 48)

Por su lado, Parodi (2010) define la LC como una metodología para tratar objetiva y estadísticamente la lengua en uso, a fin de estudiar cualquier dominio lingüístico presente en corpus compilados para tal efecto, del cual se pueden obtener datos valiosos como evidencia científica para explicar fenómenos del lenguaje, como por ejemplo, los usos de términos y abreviaciones especiales en ámbitos de la medicina (ver Moreno-Sandoval y Campillos-Llanos, 2013), la frecuencia, la dispersión y el uso de categorías léxicas (ver Dueñas Luna y González, 2012) o el análisis diacrónico y diatópico del uso del lenguaje (ver Razgovorov y Tomás, 2019).

Tanto Rojo como Parodi afirman que la LC no se trata de una metodología restrictiva, porque permite realizar estudios de la lengua desde diversas ópticas y, como enfoque metodológico, la LC posee rasgos particulares que la caracterizan. Siguiendo a Guillermo Rojo (2021), la LC es empírica, porque su interés está en la actuación de los hablantes y no en la competencia de estos, así como también se interesa en la descripción de lo que se encuentra con el uso de las lenguas y no en los universales lingüísticos.

La LC utiliza corpus textuales amplios como base del análisis y sus datos proceden de textos producidos de forma natural. Así las cosas, la LC hace uso intenso de computadoras para llevar a cabo al menos una parte del análisis. La LC emplea técnicas analíticas de carácter cuantitativo y cualitativo, como el análisis multifactorial, por ejemplo. Finalmente, la LC busca realizar análisis sistemáticos y exhaustivos de aquello que se pretende investigar en todos los casos localizados pertinentes en el corpus.

De manera que, por su carácter empírico, la LC se opone a la corriente de la lingüística racionalista (entiéndase la tendencia de la gramática generativo-transformacional chomskyana), en tanto que esta última se enfocaba en la competencia lingüística, los universales lingüísticos y en la introspección como fuente de datos (Rojo, 2021). Asimismo, la LC se opone a la lingüística descriptiva tradicional, porque la LC aspira al análisis *exhaustivo* de todos los casos pertinentes del fenómeno a estudiar y no solo a casos seleccionados. Este análisis exhaustivo es denominado *total accountability*, lo que quiere decir que la LC busca escudriñar todas las características

lingüísticas de un texto en particular, así como la forma en que estas características se interrelacionan en el contexto que lo compone (Rojo, 2021).

Con las definiciones anteriores de LC como metodología, el enfoque de la presente investigación tiene aspectos tanto cuantitativos como cualitativos, por tratarse de un estudio sobre fenómenos lingüísticos presentes en un corpus compilado, con un acercamiento epistemológico empírico basado en el análisis detallado de gran cantidad de datos. Los aspectos cualitativos tienen que ver con los análisis lingüísticos que se realizaron en cada una de las oraciones seleccionadas del corpus: el análisis de la estructura-c, de la estructura-f y de la estructura-a, y todas las tareas que se requirieron para conseguir dichos análisis. Los aspectos cuantitativos tienen que ver con los estudios estadísticos necesarios para anotar el corpus, como las listas de frecuencias de *tokens*, listas de palabras clave, y la estadística para los cálculos del acuerdo entre los anotadores, por ejemplo.

La investigación en esta tesis es de tipo básica, porque no se buscaba proponer nuevas teorías lingüísticas ni contradecir las existentes. El análisis está basado en corpus, con un enfoque de arriba hacia abajo, y de abajo hacia arriba (Völkel y Kretzschmar, 2021), porque se parte de una hipótesis con base en la jerarquía de prominencia de los roles temáticos que emplean Bresnan et al. (2016) en LFG. Es también un análisis basado en corpus porque depende de otros corpus anotados para llevarse a cabo el análisis de corpus y su anotación (Völkel y Kretzschmar, 2021), como se verá en párrafos posteriores.

Es, además, una investigación de alcance exploratorio, principalmente, porque se encontraron pocos antecedentes que ofrecieran una guía replicable. Su diseño es empírico, por tratarse de una propuesta de anotación de etiquetas lingüísticas para roles temáticos en español, y el método de análisis es por medio del análisis textual de corpus en español, empleando la perspectiva de la Lingüística de Corpus.

### ***1.8.3 Descripción del conjunto de datos***

Los datos del corpus fueron recogidos con métodos de búsqueda y recuperación de información (*Information Retrieval*, ver más sobre esta técnica en Manning, Raghavan y Schütze, 2008) para el Proyecto C0095 «Incorporación de referentes de pronombres clíticos del español en modelos de representación para tareas del procesamiento del lenguaje natural», del Instituto de Investigaciones Lingüísticas de la Universidad de Costa Rica. Los objetivos originales del *Corpus C0095* (para diferenciarlo del corpus con el que se desarrolla esta investigación), fueron los análisis

cualitativo y cuantitativo para la recuperación del sentido de pronombres en español para el proyecto anteriormente señalado. La fecha de recolección fue el 28 de julio 2020.

La temática y el volumen de datos del *Corpus C0095* lo convierten en una muestra ideal para crear un subcorpus y anotar los argumentos verbales con el análisis sintáctico, el gramatical y el semántico para los roles temáticos. Las razones que justifican esta elección se basan en que este corpus consta de datos en lenguaje natural de artículos de páginas web en español provenientes de Costa Rica, España y América Latina (ver gráfico 1). Los temas del *Corpus C0095* cubren las siguientes entradas:

- sobre COVID-19
- información COVID-19
- tratamientos COVID-19
- inyecciones COVID-19
- expertos COVID-19
- lugares COVID-19
- síntomas COVID-19
- producción COVID-19
- políticas COVID-19
- economía COVID-19
- contagio COVID-19

#### ***1.8.4 Procesamiento del Corpus COVID-19***

En este apartado se expondrá el procesamiento del *Corpus COVID-19*. En primer lugar, se indican los pasos para completar los objetivos específicos y cuáles herramientas permitieron organizar, seleccionar y anotar los textos del corpus. La investigadora decidió dividir los objetivos específicos en tareas para facilitar el orden de este apartado. De manera que, para crear un subcorpus con los 200 verbos más frecuentes; clasificar los argumentos verbales de las frases del subcorpus en cuanto a su estructura sintáctica, sus funciones temáticas y a sus funciones gramaticales, para luego elaborar el modelo de etiquetado de roles temáticos para los anotadores humanos y para *ChatGPT*, y, finalmente, calcular estadísticamente el acuerdo entre los anotadores, se necesitaron los pasos y las herramientas que se describen a continuación.

##### **1.8.4.1 Creación del Corpus COVID-19**

*Sketch Engine*<sup>58</sup> (Lexical Computing Ltd., s.f.) ofrece distintas herramientas para trabajar los corpus, como las categorías gramaticales (*Part of speech*), *N-grams*, concordancia, palabras clave, bosquejo de palabras, lemas, distribución de frecuencias, GDEX (*Good Dictionary EXamples*), tesaurus, extracción de términos y lista de palabras. Por la facilidad de conseguir todos estos análisis en una sola plataforma, la investigadora la seleccionó para llevar a cabo la

---

<sup>58</sup> Disponible en <https://www.sketchengine.eu/>. [Última consulta el 1 de junio de 2023].

preselección y la selección del *Corpus COVID-19* con los 200 verbos empleados con mayor frecuencia. *Sketch Engine* brinda una versión de prueba gratuita, pero con acceso limitado a sus aplicaciones, razón por la cual la investigadora decidió optar por la suscripción mensual de 5.99 euros para llevar a cabo esta tesis. Los costos de esta suscripción han corrido por cuenta de la tesiaría.

Para crear el *Corpus COVID-19*, se cargó el archivo del *Corpus C0095* a *Sketch Engine*, y luego se le seleccionó para trabajarlo. Con ‘*WORDLIST*’ se obtuvo la lista total de la frecuencia de los verbos en el corpus. Para depurar la lista de los 200 verbos más frecuentes, se consideró que el lema verbal no debía estar en función auxiliar, ni ser tampoco un verbo copulativo, sino que debía estar en su forma conjugada, ya fuera en oraciones activas o pasivas, en indicativo o subjuntivo (como en frases subordinadas). En los casos de tener que eliminar verbos por no contar con formas verbales válidas para su estudio, se agregaron otros verbos del corpus, según se bajaba en la frecuencia de uso.

Para cada uno de los verbos que conformarían el *Corpus COVID-19* se revisó que cumpliera con los requisitos arriba mencionados y se procedió a descargar de *Sketch Engine* un documento en formato .csv con todas las frases en donde aparecía el verbo. Se repitió el proceso hasta tener la lista de frases total con 201 verbos más frecuentes. Cada uno de los documentos .csv se abrieron en Excel y se agruparon por frecuencia, de manera que, por facilidad de trabajo, se crearon 21 documentos .xlsx con 10 hojas de cálculo cada uno. El archivo 1, por ejemplo, contiene los 10 primeros verbos más frecuentes. La investigadora procedió a limpiar cada una de las hojas de cálculo para eliminar duplicados. Con este paso se limpiaron los datos para asegurar de que no hubiera errores de etiquetado por parte de la plataforma. En columna A de la hoja de cálculo, la investigadora asignó un ID único para cada una de las frases; mientras que en la columna B se mantuvo la referencia de donde provenía la frase y en C la frase por analizar. Por razones de tiempo y de la extensión de la presente investigación, se decidió elegir cuatro frases por verbo para dar, de cierta manera, un balance al corpus total. No obstante, en algunos casos se encontraron únicamente dos o tres frases en verbos que eran de alta pertinencia por la temática del COVID-19 y se decidió incluirlos en el corpus por esta razón.

#### 1.8.4.2 Análisis de las estructuras-c, -f y -a

Es preciso adelantar aquí que, para mantener el orden y la concordancia de escritura en este apartado y en el capítulo de los resultados, cuando la investigadora se refiera a los lemas verbales lo hará siempre en mayúscula (VIVIR); en minúsculas o según sus respectivas etiquetas (SUJ, OBJ, etc.) para las funciones gramaticales; y en minúsculas y entre < > (<agente>, <instrumento>, etc.) para los roles temáticos cuando no se usen las etiquetas para estos, como lo proponen en la Gramática Léxico Funcional.

Una vez que se definió la lista completa de las frases por analizar, se creó otro documento Excel, con tres hojas de cálculo distintas. En cada hoja se creó una tabla con la misma lista definitiva de las frases del corpus. Estas tablas conformaron la hoja de trabajo para cada una de las estructuras: la estructura-c con los constituyentes; la estructura-f con las funciones gramaticales; y la estructura-a con los roles temáticos. En las listas para el análisis de cada una de las estructuras se mantuvo el mismo orden de las primeras columnas: ID, Referencia, Frase y LEMA y en las demás columnas se fueron agregando las categorías correspondientes a los análisis de cada estructura.

En la estructura-c se clasificaron los constituyentes o sintagmas del verbo de cada frase en SN, SV, SP, SAdv, SAdj, SConj y O (estas etiquetas se explicarán en el capítulo de los resultados). Se decidió clasificar únicamente los sintagmas mayores, puesto que los más detallados, como el de los determinantes no aportarían información en cuanto al tipo de rol temático en la estructura-a. Para la estructura-f, las funciones gramaticales en que se clasificaron los sintagmas de la estructura-c fueron: SUJ, V, OBJ, OBJ $\theta$ , OBL $\theta$ , XCOMP OBJ, XCOMP Prep, Pred, ADJUNTO, CONJ y CAG.

A diferencia de las dos primeras estructuras, en el análisis de la estructura-a se agregó otra columna llamada “ID Sentido”, en donde se le asignó un sentido al uso del verbo en cada una de las frases analizadas. Para clasificar los roles temáticos se siguieron los propuestos por LFG y, en parte, los del *English Propbank Annotation Guidelines* (Bonial, Hwang, Bonn, Conger, Babko-Malaya y Palmer, 2015) (Figura 18):

### Figura 18

*Etiquetas de los roles temáticos de los argumentos del verbo*

<b>ARG0:</b> agente	<b>ARG3:</b> beneficiario	<b>ARG-LOC:</b> lugar
<b>ARG01:</b> causa	<b>ARG4:</b> instrumento	<b>ARG-DIR:</b> direccional
<b>ARG1:</b> paciente	<b>ARG5:</b> experimentador	<b>ARG-GOL:</b> meta
<b>ARG2:</b> tema	<b>ARG-COM:</b> comitativo	<b>ARG-TMP:</b> tiempo

Para los adjuntos o modificadores del verbo, se establecieron las siguientes etiquetas representadas en la Figura 19:

### Figura 19

*Etiquetas de los adjuntos o modificadores verbales*

<b>ARG-EXT:</b> extensión	<b>ARG-DIS:</b> discursivo
<b>ARG-MOD:</b> modal	<b>ARG-ADV:</b> adverbial
<b>ARG-PRP:</b> finalidad	<b>ARG-NEG:</b> negación
<b>ARG-CAU:</b> causativa	<b>ARG-MNR:</b> manera
<b>ARG-REC:</b> recíproco, reflexivo, pronominal, impersonal	

Como apoyo para la anotación efectiva de las funciones gramaticales y de los roles temáticos se seleccionó el analizador lingüístico de textos *TextServer* (Padró Cicera, 2022) y *Freeling*, de la Universidad Politécnica de Cataluña, para obtener el POS de las oraciones y el análisis de roles semánticos, por medio del “Análisis de roles semánticos (deproles)”, que crea automáticamente un archivo con esta información con un formato de salida en .xml, .json, .conll o .naf, según se prefiera. La tesiaría creó una carpeta con el archivo de cada una de las frases estudiadas para futura referencia y para contrastar el etiquetado automático de estas herramientas con el llevado a cabo manualmente.

Para corroborar el etiquetado de los roles temáticos dados en cualquiera de los dos *parsers* mencionados en el párrafo anterior, la investigadora se apoyó en la plataforma del Léxico verbal del corpus *SenSem* (Grup de Recerca Interuniversitari en Aplicacions Lingüístiques, s.f.), más conocida como *GRIAL*, la cual se consideró también como un diccionario de sentidos y de definiciones, con la cual se le asignó un ID de sentido a muchos de los verbos que se analizaron en el *Corpus COVID-19*. Para los casos en donde *GRIAL* no contaba con ejemplos para los verbos del *Corpus COVID-19*, la investigadora recurrió al diccionario en línea *The Free Dictionary*, by

*FARLEX* (Farlex, s.f.), en donde se compilaron los sentidos y las definiciones para el resto de los verbos del corpus.

Para este proceso, se creó una hoja de EXCEL con la lista de los verbos seleccionados en una columna; otra con el ID del sentido verbal disponible o pertinente para los ejemplos de las frases que se iban a estudiar; otra columna con los posibles roles temáticos, particularmente los provenientes de los análisis del *GRIAL*; otra con la clase aspectual (aunque esta no se discute en la tesis) y, finalmente, una columna con la fuente de donde se tomaron los datos (*GRIAL* o *thefreedictionary.com*).

Para los casos en donde no se contó con un etiquetado automático confiable de los argumentos verbales en español por parte de *FreeLing* o *TextServer* o el *GRIAL*, se recurrieron a los principios que LFG aplica: el UTAH y el Criterio  $\theta$ . Asimismo, se tomaron en cuenta los rasgos  $[\pm o]$   $[\pm r]$  de los *Mapping Principles* y el *Wellformedness Condition*. Con estas plataformas, *TextServer*, *FreeLing*, *GRIAL*, *thefreedictionary.com* y con la teoría LFG se garantiza haber asignado manualmente etiquetas justificadas desde diversas perspectivas lingüísticas para los roles temáticos encontrados en el *Corpus COVID-19*.

#### 1.8.4.3 Desarrollo del modelo de etiquetado de los roles temáticos

Una vez que la investigadora analizó las tres estructuras en su totalidad para las 801 frases, procedió a crear un modelo de anotación para anotadores humanos, con el cual replicar la clasificación de los roles temáticos sobre el mismo corpus estudiado. Lo primero que fue necesario ajustar fue la cantidad de etiquetas que se darían a los anotadores para que trabajaran el modelo. De las 21 etiquetas que la investigadora anotó en el corpus original, se redujo el número a solamente 8 posibles roles, dejando únicamente los argumentos en el estricto sentido de su definición (ver marco teórico) y los adjuntos de tiempo y de lugar únicamente. La razón detrás de esta disposición fue porque, de analizarse las 21 etiquetas, se crearía un alto nivel de confusión y ambigüedad en una tarea que, ya de por sí, es difícil de concordar.

En adición a excluir las etiquetas de los modificadores y su análisis, también se redujo el número de roles temáticos para los argumentos del verbo. En este punto se tomó la difícil decisión de combinar roles que compartieran ciertas características para ofrecer todavía menos opciones de selección a los anotadores. De manera que el modelo quedó con los siguientes roles temáticos: agente, tema, meta, experimentador, fuente, instrumento, lugar y tiempo. Una vez establecidas

estas etiquetas, se decidió que la mejor manera de evaluar la anotación de los argumentos en cada frase sería a través de una encuesta, en donde se le presentara a los anotadores una lista de frases con uno de sus argumentos subrayados para que lo clasificaran tomando en cuenta únicamente los ocho roles temáticos arriba señalados. Con esto en mente, se creó otro documento de Excel, se tomó la lista del análisis de la estructura-a, de la cual se mantuvo el ID original, el LEMA y la frase con el argumento subrayado. Luego, todos los argumentos y los adjuntos de las frases que correspondieran con los ocho roles temáticos que la investigadora definió, fueron resaltados con subrayado, en negrita y en color rojo, en cada una de las 801 oraciones del *Corpus COVID-19*.

Luego se clasificaron las frases por tipo de rol temático; se corrigió cualquier error de etiquetado del corpus original y, finalmente, se creó una lista con todos los argumentos subrayados, la cual quedó conformada por 1531 ítems. Cabe mencionar que, en muchos casos, una misma frase aparecía en la lista con distintos argumentos marcados, algunas veces con el argumento para su sujeto, o para su objeto directo, o para su argumento de tiempo, etc. Por esta razón, se optó por agregar otro ID para distinguir y hacer única cada entrada (frase) del modelo. Este nuevo número con el nombre de “PK” se asignó luego de haber aplicado la fórmula de Excel RAND (ALEATORIO en español), para que a la hora de asignar las frases a cada anotador estas no fueran consecutivas de un mismo verbo o de un mismo rol temático.

Originalmente, se tomaron 15 estudiantes del Programa del Posgrado de Maestría Académica en lingüística, específicamente del curso Análisis sintáctico (SP-7184), impartido por el profesor Jorge Antonio Leoni de León, para que fueran los anotadores del modelo. De manera que el modelo se formuló para que cada frase fuera evaluada por al menos 3 anotadores distintos y para que cada anotador no evaluara más de una vez cada ítem. Para conseguir esto, los 15 anotadores humanos se dividieron en 5 grupos, con 3 anotadores cada uno, para anotar, aproximadamente, 306.2 ítems cada persona. Después se tomaron las 1531 frases del modelo y se creó otra lista con tres copias consecutivas de las 1531 frases, para conseguir un total de 4593 ítems. Las frases por etiquetar se asignaron a los anotadores por orden del “PK”: las primeras 305 frases para el anotador 1, las 306 siguientes para el anotador 2 y así, consecutivamente.

La investigadora de esta tesis procedió a crear 15 carpetas con los nombres de “Anotador 1”, “Anotador 2, ... y “Anotador 15”, las cuales contenían la lista de  $\pm 306$  frases en un documento Excel para cada anotador y la guía de anotación en un documento Word que explicaba los pasos a seguir en la encuesta y la definición de los ocho roles temáticos que debían emplear para tal tarea.

Esta guía se puede encontrar en el Anexo A y la encuesta en el Anexo B. Estas carpetas se compartieron con el profesor Leoni de León para que él las asignara a cada anotador, de manera que la tesiaría no tuviera conocimiento de quienes anotaron qué.

Para que *ChatGPT* fuera partícipe también del proceso de anotación de la encuesta, como se propuso en los objetivos, se tuvieron que hacer transformaciones al modelo de anotación. En primer lugar, la lista de las 1531 frases con los argumentos subrayados, en rojo y en negrita tuvieron que modificarse, porque *ChatGPT* no reconocía estas marcas de formato en los *prompts* que se aplicaban como interrogantes. *ChatGPT* informó a la investigadora que el formato que reconocería debería estar entre asteriscos (\*...\*). Así las cosas, la tesiaría procedió a cambiar las marcas dadas al modelo para humanos para que cada argumento marcado quedara entonces entre asteriscos. Por ejemplo, la frase para el lema AFIRMAR pasó de tener la forma en (17) a quedar como en (18):

- (17) «**El infectólogo Carlos Álvarez** afirmó que ante síntomas gripales lo mejor es aislarse ya que podría ser COVID-19».
- (18) «**\*El infectólogo Carlos Álvarez\*** afirmó que ante síntomas gripales lo mejor es aislarse ya que podría ser COVID-19».

Una vez que el modelo estuvo listo, se procedió a entrevistar a *ChatGPT*. La investigadora consideró necesario suscribirse a la versión pagada de este modelo LLM para tener acceso a *GPT-4* y estos gastos corrieron por cuenta de la tesiaría. A pesar de la suscripción, *GPT-4* tiene un límite de 25 mensajes o *prompts* cada 3 horas, por lo que la mayoría de la entrevista tuvo que llevarse a cabo con el modelo anterior, *GPT-3.5*, porque no siempre la entrevista podía realizarse cada tres horas, puesto que esto habría extendido en tiempo esta tarea.

El *prompt* empleado fue: «Quiero saber cuál es el rol temático del segmento entre asteriscos de las frases que te voy a ir dando. Para esto quiero que uses únicamente el conjunto de roles temáticos: agente, quien realiza la acción; tema, a quien le ocurre la acción; meta, el lugar o persona hacia donde se dirige un movimiento o acción; experimentador, quien percibe o experimenta sensaciones provocadas por una fuente; instrumento, con qué o con quién se realiza una acción; fuente, provoca sensaciones en otra entidad o es el origen de donde parte una acción; lugar, donde se desarrolla la acción; y tiempo, el período relacionado con el desarrollo de la acción». La tesiaría fue preguntando, una a una, las 1531 frases al modelo, y fue marcando las respuestas dadas en el documento Excel con el modelo de anotación para *ChatGPT*.

Es importante indicar aquí que la entrevista a *ChatGPT-3.5* y *ChatGPT-4* se llevó a cabo con *zero-shot learning*, puesto que no se buscaba entrenar al programa, sino obtener respuestas prontas, similares a las humanas. Los anotadores humanos tampoco fueron entrenados en el análisis de roles temáticos, sino que trabajaron con la misma guía que se le dio a *ChatGPT*. Además, no hubo pruebas con cambios de temperatura, porque se prefirió trabajar con los ajustes predeterminados por los desarrolladores. La única variable que se modificó en el chat fue el tono, para que este fuera siempre profesional. Los resultados y las dificultades de esta tarea serán discutidas en el capítulo de desarrollo.

#### 1.8.4.4 Evaluación del acuerdo de los anotadores

El acuerdo entre anotadores (*IAA*, por sus siglas en inglés) es un método para medir la concordancia o consistencia entre dos o más codificadores quienes, individualmente, anotaron o clasificaron los mismos datos de un corpus (ChatGPT, 2023). La razón por la cual se requiere este paso en la presente investigación es porque es necesario evaluar la calidad de las etiquetas de los roles temáticos que se asignaron, tanto por humanos, como por *ChatGPT*, al *Corpus COVID-19*.

Existen varias métricas para medir el *IAA*, y entre ellas las más comunes en LComp y LC son los coeficientes kappa de Cohen y el alfa de Krippendorff (empleado en el estudio de Reiss, 2023), así como el de kappa de Fleiss (Pustejovsky y Stubbs, 2013). El coeficiente de kappa de Cohen mide el acuerdo entre dos anotadores tomando en cuenta la posibilidad del acuerdo que ocurre por azar. Por su lado, el coeficiente alfa de Krippendorff mide el desacuerdo entre los anotadores y requiere de métodos de *bootstrap* o de muestra de arranque (Benites, 2022), que es «un tipo de remuestreo en el que se extraen repetidamente grandes cantidades de muestras más pequeñas del mismo tamaño, con reemplazo, a partir de una única muestra original» (Benites, 2021<sup>a</sup>, párr. 1) y calcularlo es computacionalmente complejo y sus resultados son comparables con los del coeficiente kappa de Cohen, por eso se prefiere menos (Benites, 2022).

Pustejovsky y Stubbs (2013) define el coeficiente kappa de Fleiss es una medida estadística para calcular el acuerdo real y el acuerdo esperado entre los anotadores, teniendo en cuenta el juicio aleatorio o la posibilidad de que dos o más anotadores hayan dado la misma respuesta al azar en una encuesta, y la fórmula para calcular este valor es la siguiente:

**Figura 20**

*Fórmula del coeficiente kappa de Fleiss*

$$k = \frac{P - P_e}{1 - P_e}$$

El coeficiente kappa de Fleiss se prefiere cuando diferentes grupos de anotadores califican diferentes objetivos (Laerd Statistics, 2019), lo cual concuerda con la presente investigación. Laerd Statistics (2019) propone tres suposiciones o requerimientos para emplear efectivamente esta métrica y, basadas en ellas, la investigadora de esta tesis justifica calcular el *IAA* por medio de este coeficiente:

- Las variables de las respuestas se miden en categorías y estas son mutuamente excluyentes: las categorías son las etiquetas de los ocho roles temáticos y que los anotadores deben asignar uno por frase únicamente.
- Todos los anotadores tienen las mismas categorías para analizar las frases: todos los anotadores siguen la misma guía y clasifican los argumentos según los mismos ocho roles temáticos.
- Los anotadores deben ser independientes, de manera que el juicio de uno no afecte el de otros anotadores.

Estas suposiciones ofrecieron a la investigadora razones suficientes para confiar en los cálculos del acuerdo entre los anotadores, tanto entre humanos, como para comparar los resultados de *ChatGPT*. Los resultados se presentarán y explicarán en el siguiente capítulo, en donde se discutirán los problemas y los obstáculos encontrados durante el proceso de investigación con las herramientas empleadas en este apartado. A continuación, se presenta el proceso para calcular  $\kappa$ .

La fórmula de la Figura 20, explicada en palabras, significa que el coeficiente kappa toma en cuenta los siguientes valores:

$$(16) \quad (\kappa) = \frac{\text{acuerdo observado entre los anotadores} - \text{acuerdo esperado entre los anotadores}}{1 - \text{acuerdo esperado entre los anotadores}}$$

En (16), el acuerdo observado de los anotadores ( $P$ ) corresponde con la exactitud (*accuracy*) con la que estos respondieron la encuesta y se calcula por medio de la siguiente fórmula:

**Figura 21**

*Fórmula para calcular el acuerdo observado entre los anotadores*

$$P_i = \frac{(\sum_{c=1}^k a_{ic}^2) - (a)}{a(a - 1)}$$

En la Figura 21,  $P_i$  se calcula para cada fila sumando los valores al cuadrado de las categorías (los roles temáticos) en cada columna y moderando la salida por el número total de anotaciones por cada anotador (Pustejovsky y Stubbs, 2013).

Para obtener el acuerdo esperado entre los anotadores ( $P_e$ ), se calcula  $P_c$ , en donde  $c$  equivale a cada una de las categorías (los roles temáticos agente, tema, meta, etc.), sumando los valores en su columna dividida por el número de anotadores multiplicado por el número de anotaciones que creó cada anotador (Pustejovsky y Stubbs, 2013). En la Figura 22 se muestra la fórmula, en donde la segunda parte de esta indica que la suma de todas las anotaciones que un anotador hizo y dividida entre el número de anotaciones que cada anotador hizo individualmente, da un valor igual a 1.

**Figura 22**

*Fórmula para calcular el acuerdo esperado entre los anotadores*

$$P_c = \frac{1}{Aa} \sum_{i=1}^A a_{ic}, \quad 1 = \frac{1}{a} \sum_{c=1}^k a_{ic}$$

Cuando se tienen los valores de  $P_i$  y de  $P_c$  se procede a calcular el valor del coeficiente de kappa de Fleiss o IAA de la Figura 19. Landis y Koch (1977) (citado en Pustejovsky y Stubbs, 2013; pág. 131) propuso las directrices de la Tabla 4 para interpretar los resultados de  $\kappa$ , el cual puede ir entre 0 a 1, en donde 0 se considera que no hubo acuerdo y 1 equivale a un acuerdo perfecto. En algunos casos se pueden tener valores negativos, los cuales indican que los valores fueron menores a los esperados por el acuerdo casualidad (Benites, 2021b).

**Tabla 4***Directriz para interpretar los valores de  $\kappa$* 

$\kappa$	Nivel de acuerdo
< 0	pobre
0.01-0.20	ligero
0.21-0.40	justo
0.41-0.60	moderado
0.61-0.80	sustancial
0.81-1.00	perfecto

Para calcular el coeficiente de kappa de Fleiss en esta investigación, se tomaron las respuestas de los anotadores y se agruparon de acuerdo con el grupo que les correspondía. Luego se les asignó una etiqueta con el número de grupo al que estas pertenecían, para poder clasificarlas luego. Las respuestas de los cinco grupos se unificaron en una sola lista para calcular cuántas respuestas por PK (o argumento) se obtuvieron entre todos los anotadores. Para esta tarea, se recurrió a un *script* del programa *SQL Server*<sup>59</sup>, un sistema de gestión de bases de datos relacionales (*Relational Database Management System, RDBMS*) en el lenguaje *Transact-SQL (T-SQL)*. Este *script* fue ejecutado por un administrador de bases de datos profesional<sup>60</sup> y con él se extrajeron también las respuestas nulas que son aquellas que carecían de etiqueta o tenían más de una respuesta por un mismo anotador, también los casos de falsas respuestas a causa de caracteres de espacios en blanco (*whitespace characters*) dejados por los anotadores. Las respuestas falsas se corrigieron en los datos para obtener el conteo real y las nulas se descartaron en el cálculo del *IAA*.

La lista con el conteo de las respuestas válidas se clasificó por el número de grupo para calcular el coeficiente de  $\kappa$  individualmente. Dentro de cada grupo se subdividió la lista por el número total de anotadores que dieron una etiqueta a cada frase, porque en el valor de  $\kappa$  se consideran el número de anotadores y el número de frases anotadas. Se descartaron los casos en donde solo apareció un anotador por frase. Para los casos en donde no hubo acuerdo entre los anotadores, se empleó un anotador ajeno al proceso de anotación anterior, para que definiera una respuesta final y así quedarán todas las frases con una etiqueta para el argumento en cuestión. Esta lista final con las etiquetas humanas se comparó luego con la lista original con las anotaciones

<sup>59</sup> Disponible en <https://www.microsoft.com/es-es/sql-server/sql-server-downloads>. [Última visita el 6 de junio de 2023].

<sup>60</sup> El ingeniero en sistemas informáticos, Carl Albert Florin, quien tiene amplia experiencia en administración de datos y en *SQL Server*.

manuales de la investigadora. Se usó nuevamente otro *script* de *T-SQL Server* para obtener el conteo de anotaciones por PK y calcular con este conteo el coeficiente de  $\kappa$  de Fleiss.

Por medio de otro *script* de *SQL Server*, se combinaron las respuestas de la encuesta de *ChatGPT* con las de los anotadores humanos. Se siguió el mismo procedimiento para descartar los casos con respuestas nulas (humanas y de AI) y la lista combinada se clasificó luego por número de grupo y dentro de cada grupo, por número de anotadores, para calcular el *IAA*. Finalmente, se comparó el desempeño de *ChatGPT* con respecto a las anotaciones manuales de la investigadora, también con la ayuda de *SQL Server*, para obtener el valor de  $\kappa$ .

En resumen, en este último apartado del capítulo I se han explicado los pasos que se siguieron para crear una aproximación lingüística en el diseño de un corpus anotado en español sobre COVID-19 para sistemas de pregunta-respuesta. Se han definido las herramientas que fueron necesarias para obtener resultados precisos y respaldados con bases teóricas lingüísticas y estadísticas. Se ha justificado la perspectiva lingüística, el enfoque, el tipo de investigación, el alcance, el acercamiento epistemológico, el método de análisis y el diseño de este. En el capítulo II se reportarán los resultados y su correspondiente análisis, así como las dificultades encontradas en los pasos y en las herramientas empleadas.

## CAPÍTULO II. DESARROLLO

### 2.1 Presentación y análisis de los resultados

En este segundo capítulo se darán a conocer los resultados del objetivo general propuesto para esta investigación: desarrollar una propuesta de modelo de anotación de roles temáticos en español específico para sistemas de pregunta-respuesta. En el apartado [2.2](#) se presentarán los resultados y su análisis de acuerdo con el orden de los objetivos específicos que se propusieron, con la idea de seguir la lógica con la cual se llevó a cabo la investigación. En [2.3](#) se propondrán las discusiones de estos resultados.

#### 2.1.1 Crear un subcorpus con los 200 verbos más frecuentes

En el apartado [1.8.3](#) se estableció que el *Corpus COVID-19* es un subcorpus o corpus secundario del *Corpus C0095* que, por su contenido y tamaño, se decidió que era el ideal para trabajar en esta investigación. El *Corpus C0095* tiene el formato de documento de texto (\*.txt), con la codificación UTF-8 y el salto de línea en Unix (LF). Este documento se subió a la plataforma *Sketch Engine*, para analizar sus datos y para seleccionar un subcorpus a partir de él. Esta es la misma herramienta con que López Rodríguez (2020) clasificó los verbos en textos médicos, como se mencionó en los antecedentes. Se encontró que este archivo lo componen una totalidad de 57 documentos, de los cuales dos páginas de Wikipedia conforman la mayoría de ellos (ver la Figura 23), y de los que se compilan 89 685 *tokens*, 76 015 palabras y 3 761 oraciones (ver la Tabla 5).

**Tabla 5**

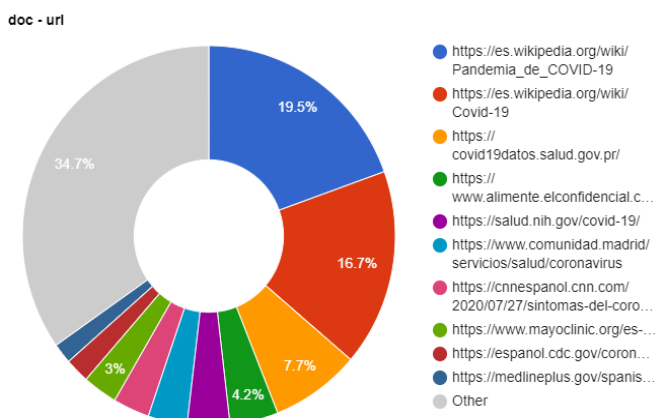
*Datos del Corpus C0095 tomados de Sketch Engine*

#### COUNTS

<b>Tokens</b>	89,685
<b>Words</b>	76,015
<b>Sentences</b>	3,761
<b>Paragraphs</b>	2,446
<b>Documents</b>	57

**Figura 23**

*División de los URL de los documentos del Corpus C0095*



El siguiente paso fue encontrar los verbos más frecuentes del corpus con ‘Wordlist’, como se explicó en 1.8.1. En la Tabla 6 se ofrece una muestra de la lista de los verbos más frecuentes, según lo que *Sketch Engine* recuperó, en donde en la primera columna se encuentra el verbo; en la segunda, su frecuencia absoluta; en la tercera, en cuántos documentos fue frecuente (*DOCF*); en la cuarta, la frecuencia reducida media (*ARF*); y, en la quinta, la frecuencia media de distancia logarítmica (*ALDF*). De estos valores, solamente importó el de la frecuencia absoluta para los fines de esta tesis.

**Tabla 6**

*Muestra de la lista de frecuencias verbales en el Corpus COVID-19*

WORDLIST					Corpus Covid				
verb (938 items   11,495 total frequency)									
Verb	Frequency ? ↓	DOCF ?	ARF ?	ALDF ?	Verb	Frequency ? ↓	DOCF ?	ARF ?	ALDF ?
1 ser	1,135	57	688.18	677.27 ...	251 despedir	7	5	4.12	4.66
2 haber	673	53	374.93	361.78 ...	252 originar	7	5	4.76	5.22
3 poder	581	54	328.29	312.36 ...	253 adquirir	7	6	4.94	5.65
4 estar	393	50	208.21	195.70 ...	254 certificar	7	4	3.69	3.70
5 tener	353	46	187.14	172.19 ...	255 aprender	7	6	2.27	2.03
6 deber	211	41	109.98	102.00 ...	256 encargar	7	4	4.01	4.21
7 presentar	158	38	74.04	66.76 ...	257 quitar	7	4	3.38	3.21
8 hacer	140	43	78.33	76.04 ...	258 evidenciar	7	3	3.29	3.44
9 infectar	109	28	52.46	46.20 ...	259 abrir	7	4	2.78	2.68
10 incluir	103	29	54.55	56.14 ...	260 adoptar	7	5	4.45	4.87
11 unir	100	28	52.25	46.02 ...	261 mencionar	7	4	4.70	5.02
12 usar	99	20	33.75	26.54 ...	262 solicitar	7	7	5.95	6.63
13 dar	98	31	56.26	59.17 ...	263 conseguir	7	7	4.07	4.08
14 vacunar	95	23	29.47	23.63 ...	264 expulsar	7	5	4.30	3.99
15 decir	94	28	43.10	38.53 ...	265 especificar	7	3	2.56	2.52

En la Tabla 6, se encuentra que la totalidad de palabras etiquetadas como ‘verbo’ es de 938 ítems; y en donde los verbos más frecuentes son los esperados auxiliares y modales. Cabe señalar aquí que *Sketch Engine* aplica su propio *parser* para conseguir estas listas; en consecuencia muchos ítems fueron etiquetados erróneamente como verbos cuando no lo eran, como en ‘*d’habitatge*’: «<s>El Sindicat d’Habitatge de Nou Barris cree que se están acumulando los desalojos en previsión de que acabe la moratoria.</s>»; o ‘*been*’, de una oración que estaba en inglés: «<s>When you’ve been fully vaccinated.</s>» y que la plataforma analizó como si esta fuese una frase en español.

*Sketch Engine* clasificó también como verbo todos los lemas que encontró como tales, por lo que fue normal leer frases como «Consultado el 19 de abril de 2020», en donde se esperaba el verbo CONSULTAR (con frecuencia 84) en su forma conjugada y no en función de participio. En la selección de la lista para el *Corpus COVID-19*, la investigadora encontró verbos frecuentes con este comportamiento, o sea, con listas de oraciones con pocos o ningún ejemplo en donde el lema estuviera en la forma verbal conjugada esperada. Estos verbos se descartaron y se continuó buscando otros más propicios. Algunos de los verbos que reportaron este problema fueron HOSPITALIZAR, RELACIONAR, PARTIR (aparecía como “a partir” en las frases extraídas), AISLAR, ACUMULAR, ALEJAR, PESAR (aparecía como “a pesar”), MODERAR, ACERCAR, ILUSTRAR, ACABAR, EXPONER, MOVER, ACERCAR, CONTAMINAR, ELEVAR, PRIVAR, ARCHIVAR, MEDIAR, MEDICINAR, EMBARAZAR, entre otros.

Cuando se tuvo la lista con los posibles 200+ verbos principales, se procedió a crear archivos por verbo con la lista de todas sus frases en Excel, como se explicó en la metodología. Estas listas se limpiaron de elementos no deseados, como caracteres del tipo “<s>”, “</s>” en (17), (18) y (19), números de referencia a fuentes como “[233]” en (18) y (19), o frases que no pertenecían a la oración del verbo conjugado en cuestión, como en (19), que, por falta de espacios o puntuación en el texto original, *Sketch Engine* la recuperó como una sola frase (se subraya el verbo que se buscaba analizar):

- (17) «<s> El menor de 15 años manifestó que sí se quiere vacunar alegando que si se suministra "es porque es confiable". </s>»
- (18) «<s> Para el 6 de enero de 2022, se habían administrado 9 370 millones de dosis de las vacunas contra la COVID-19, distribuidas solamente entre el 59% de la población mundial.[233] </s>»

- (19) «un derivado menos tóxico de cloroquina, sería más potente para inhibir la infección por SARS-CoV-2 in vitro.[280][281] El 16 de marzo de 2020, una importante autoridad francesa y asesor del Gobierno francés sobre COVID-19, el profesor Didier Raoult del Instituto Universitario Hospitalario de Enfermedades Infecciosas (IHU-Méditerranée infection) en Marsella (Bouches-du-Rhône, Provenza-Alpes-Côte d 'Azur), anunció que un ensayo con 24 pacientes del sureste de Francia había demostrado que la cloroquina es un tratamiento efectivo para COVID-19.[282][283] Se administraron 600 mg de hidroxicloroquina (marca Plaquenil) a estos pacientes todos los días durante 10 días. </s>»

Se eliminaron también los duplicados de frases y se asignó un número de ID a todas las frases. Para mantener un registro y corregir posibles errores, la investigadora creó una tabla en una hoja de cálculo en Excel con el verbo y con el rango de ID correspondiente para su lista de frases (ver la Tabla 7), la cual sirvió para corregir un error manual que se había cometido y que luego resultó en que dos frases con verbos distintos tuvieran el mismo ID.

**Tabla 7**

*Registro de ID para archivos del subcorpus COVID-19*

# Archivo	VERBO	ID
1	HABER	1-75
1	PRESENTAR	76-211
1	INFECTAR	212-313
1	INCLUIR	314-411
1	UNIR	412-500
1	USAR	501-593
1	DAR	594-683
1	CAUSAR	684-772
1	AFECTAR	773-851
1	MOSTRAR	852-921

**Tabla 8**

*Lista de frases por verbo descargada de Sketch Engine*

A	B	C
1	academic use only	
2	method name: concordance	
3	corpus: user/Yislen/corpus_covid	
4	subcorpus: -DAR	
5	concordance size: 98	
6	query: Query:[lempos=="dar-v"]	
ID	Reference	Sentence
8	594 medlineplus.gov	<s> Este le dará las guías locales más recientes acerca del examen. </s>
9	595 medlineplus.gov	<s> NO le dé aspirin a un niño (menor de 18 años) a menos que el proveedor de su hijo se lo indique. </s>
10	596 sabervirtv.com	<s> Esos casos se dan sobre todo en pacientes ingresados, medicalizados. </s>
11	597 sabervirtv.com	<s> Las náuseas y vómitos también se dan en pocas ocasiones entre los enfermos de la COVID-19 (6%). </s>
12	598 sabervirtv.com	<s> Es cierto que en el COVID-19 también se da fatiga y no se puede descartar. </s>

En la Tabla 8 se muestra una parte de la lista de las frases por verbo que se descargaron de *Sketch Engine* para luego seleccionar las que se usarían en el análisis del corpus, en donde se le asignó el número de ID en la columna A, el nombre del verbo, el nombre del archivo y del corpus. El color rojo en la frase de la Tabla 8 indica que era una posible opción para las cuatro que se seleccionarían. Es importante mencionar aquí que no siempre fue posible encontrar cuatro frases para todos los verbos, porque muchas de las oraciones recuperadas no cumplían con tener el verbo conjugado en forma activa o pasiva. Hubo verbos que, por su relevancia con el tema del virus COVID-19, se mantuvieron en el corpus a pesar de tener solo tres ejemplos, como DETERMINAR, EXCLUIR, RESPIRAR, SUCEDER y ATENDER; con o solo dos ejemplos, como ELEVAR (20-21):

- (20) «Para que la alarma saltara, la frecuencia cardíaca debía e elevarse durante un espacio prolongado de tiempo, por lo que si alguno salía a correr no saltaba.»
- (21) «Además, las circunstancias actuales\_dificultan el acceso a teléfonos, líneas de atención y servicios públicos como la policía, la justicia y los servicios sociales, por lo que e elevan la vulnerabilidad de las mujeres a la violencia en el ámbito privado en todo el mundo.»

Una vez que se tuvo la lista con todas las frases por verbo, se procedió a clasificar los argumentos de todos los verbos seleccionados realizando los tres análisis propuestos por LFG en el marco teórico, como se presentarán los resultados en detalle en el siguiente apartado. La lista de los verbos estudiados se encuentra en el Anexo C.

### ***2.1.2 Clasificar los argumentos verbales en cuanto a sus funciones temáticas y a sus funciones gramaticales***

Es importante indicar que el análisis de las tres estructuras (-c, -f y -a) se llevó a cabo a la misma vez para cada una de las frases, comenzando con la estructura-c, con el orden de los constituyentes, como se explicó en el capítulo anterior; luego se analizó la estructura-f, con las funciones gramaticales de los constituyentes y, finalmente, se anotó la estructura-a, con los sentidos verbales y los roles temáticos de los argumentos. De esta manera se mantuvo el hilo conector con la relación de los análisis, en todos sus niveles, conforme se iba etiquetando cada frase. Antes de proceder con la clasificación de los argumentos y de los adjuntos modificadores con las funciones gramaticales y las semánticas fue necesario construir un diccionario con las

definiciones y los sentidos verbales con el que se pudiera asignar un sentido válido a los verbos de las 801 frases del corpus y así discernir cuáles participantes estaban involucrados en cada uno de estos eventos verbales. A continuación, se presentan los resultados de esta tarea.

### 2.1.2.1 TextServer, FreeLing y GRIAL

La idea detrás de crear una ontología o diccionario con entradas específicas de sentidos y definiciones era facilitar el etiquetado manual de los análisis de las estructuras-c, -f y -a por parte de la investigadora. Para crear esta ontología se emplearon dos herramientas: el *GRIAL* y *The Free Dictionary, by FARLEX*<sup>61</sup>, como se explicó en la metodología. El *GRIAL* (Figura 24) es la plataforma que resultó del trabajo conjunto de los estudios de Vázquez et al. (2006); Castellón et al. (2012) y Gil-Vallejo et al. (2015), todos expuestos en el estado de la cuestión de esta tesis, y que cuenta con los análisis de las funciones gramaticales, los roles temáticos, los sentidos y las definiciones para los verbos del corpus *SenSem*. El *GRIAL* se consideró como una primera base para crear el diccionario de definiciones y sentidos, con el cual se asignarían los sentidos a muchos de los verbos del *Corpus COVID-19* en la estructura-a y se contrastarían los roles temáticos asignados por la tesiaría. *PropBank* (Bonial et al., 2015), Martí Antonín et al. (2007) y García Marchena (2009) también utilizaron un “*sense ID*” para justificar el etiquetado de los roles temáticos de los argumentos verbales, fuera este manual o automático.

#### Figura 24

Plataforma del Léxico verbal del corpus *SenSem* (*GRIAL*)



The screenshot shows the GRIAL web interface. At the top, there is a navigation bar with links: Inicio, Documentación, Corpus, Léxico, Subcategorizaciones, Descargas. Below this is the GRIAL logo and the text 'Léxico verbal del corpus SenSem'. The main content area is divided into two sections. On the left, there is a search form titled 'Criterio de búsqueda' with a 'Buscar por verbo' section. The 'Idioma' is set to 'Español' and the 'Verbo' is 'defender'. A 'Buscar' button is present. Below this is a link to 'Buscar por criterios comunes'. On the right, there is a table titled 'defender' with the following data:

Sentido	Definición	Roles	Clase aspectual	Frecuencia
1	<a href="#">Dar amparo físico, moral, ideológico o legal a alguien o algo.</a>	Agente, Tema, Origen	Proceso	50/122
2	<a href="#">Demostrar cierta habilidad en alguna actividad.</a>	Agente, Manera, Tema	Proceso	1/122
3	<a href="#">Apoyar una idea o causa.</a>	Agente - Origen, Tema (objeto desplazado), Medio	Proceso	70/122
4	<a href="#">En ciertos deportes o juegos, mostrar resistencia ante el ataque del adversario.</a>	Agente	Proceso	1/122

GRIAL - Grup de Recerca Interuniversitari en Aplicacions Lingüístiques  
*SenSem* - Sentence Semantics: Creación de una Base de Datos de Semántica Oracional

La segunda herramienta ampliamente empleada fue *The Free Dictionary*, con la que se compilaron las definiciones y los sentidos verbales que no aparecían analizados en el *Corpus*

<sup>61</sup> Disponible en <https://es.thefreedictionary.com>. [Última visita el 09 de junio de 2023].

*SenSem*; con esta herramienta, además, se confirmaron los sentidos dados por *GRIAL*. Este diccionario en línea compila las definiciones de varios diccionarios reconocidos, con lo que la fiabilidad de sus resultados es más alta que el haber empleado un solo diccionario como referencia. *The Free Dictionary* no cuenta con análisis gramaticales ni semánticos; no obstante, fue un gran soporte a la hora de decidir qué número de valencia tenían los verbos del *Corpus COVID-19* y para poder clasificarlos en transitivos, intransitivos, pronominales, etc.

A pesar de que el Léxico verbal del corpus *SenSem GRIAL* está bien creado y es fruto de años de investigación, tiene varias inconsistencias que son difíciles de seguir como guía de anotación de roles temáticos y de sentidos verbales. Una de las inconsistencias encontrada fue cuando asignaron la definición para la forma pronominal a un verbo cuando la entrada al diccionario estaba en forma transitiva (Figura 25). Aunque el verbo de la Figura 25 no se analizó en el *Corpus COVID-19*, este error se encontró mientras se analizaban otros verbos relacionados con el sentido.

### Figura 25

*Inconsistencias en el Léxico verbal del corpus SenSem (GRIAL)*

aprovechar				
Sentido	Definición	Roles	Clase aspectual	Frecuencia
1	<a href="#">Servirse de algo para un determinado fin.</a>	Agente, Tema, Finalidad	Evento	109/120
2	<a href="#">Beneficiarse de alguien o algo para sacarle un provecho personal.</a>	Agente, Instrumento, Finalidad	Proceso	11/120

Otro problema del *GRIAL* fueron las etiquetas asignadas para algunos de los roles temáticos. En ninguno de los trabajos en los antecedentes (Vázquez et al., 2006; Castellón et al., 2012; Gil-Vallejo et al., 2015) se hace referencia a qué guías emplearon para decidir los rasgos de los roles temáticos, particularmente para los sujetos. En varias ocasiones quedó confuso qué tipo de rol habían asignado, como en dos frases para DAR (Figura 26), en donde el sujeto tenía los roles <agente-origen>, cuando en la teoría se prescribe que se debe asignar únicamente un rol por función gramatical y que, de haber más de un posible rol temático para ese verbo, se debe aclarar que dependerá del sentido verbal que tenga esa frase en particular y se debe etiquetar separadamente.

**Figura 26***Inconsistencias en el etiquetado del GRIAL*

dentro del jueves y que tampoco el nuevo titular de Economía , Pedro Solbes , **había dado** razón alguna .

Evento Perfectivo		
Modalidad asertiva - Polaridad positiva		
Topicalización sujeto lógico		
Sujeto	V	Objeto directo
Sintagma nominal (nombre común)		Sintagma nominal (nombre común)
Agente - Origen		Tema (objeto desplazado)

---

30 episodios de 13 minutos en el que los niños **darán** su opinión sobre , por ejemplo , qué es la felicidad y adónde van las

No factual		
Evento		
Modalidad asertiva-futuro - Polaridad positiva		
Topicalización sujeto lógico		
Sujeto	V	Objeto directo
Sintagma nominal (nombre común)		Sintagma nominal (nombre común)
Agente - Origen		Tema (objeto desplazado)

En la metodología se explicó que se creó la ontología en una hoja de cálculo con la lista de los 201 verbos del *Corpus COVID-19*, a la que se le asignaron los sentidos y las definiciones a cada uno de los verbos y, dependiendo de si estos fueron tomados del *GRIAL* o de *The Free Dictionary*. Con este material listo, se procedió a obtener las etiquetas automatizadas para la estructura-c, la estructura-f y la estructura-a, con la herramienta *TextServer*, a manera de tener otra guía de anotación para este corpus. El uso de esta plataforma es gratuito y cuando se trabaja con corpus extensos, se sugiere al usuario crear una cuenta con la que se le otorga al investigador acceso ilimitado a los servicios y a los servidores del programa. Pero, a pesar de esto, los servidores para analizar los roles temáticos no siempre estuvieron disponibles o estuvieron en mantenimiento en el momento en que se necesitaban los resultados, por lo que, en no pocas ocasiones, se empleó *FreeLing* como opción de respaldo.

Con *TextServer* o *FreeLing* se crearon documentos en formato *.text* con el análisis morfológico, el POS *tagging*, la clasificación de entidades, la desambiguación de sentidos, las dependencias sintácticas y los roles temáticos. Al emplear el “Análisis de roles semánticos (deproles)” en *TextServer*, se obtuvieron, casi siempre, los mismos resultados que cuando se empleó el “*Semantic Graph*” de *FreeLing*, la herramienta más empleada en los estudios

presentados en los antecedentes. En la Figura 27 se muestra un ejemplo del archivo con este tipo de análisis automatizado para el verbo ACABAR:

### Figura 27

*Archivo para el verbo ACABAR con el análisis deproles en TextServer*

```

## cputime=0.111892 wordcount=14
1 Algunos alguno DI0MP0 DI - - - - 2 spec - - -
2 virus virus NCMN000 NC - - 01328702-n - 5 suj - - - A0 -
3 aéreos aéreo AQ0MP00 AQ - - 01380267-a - 2 s.a - - -
4 pueden poder VMIP3P0 VMI - - 02402825-v - 5 v - dismiss.01|fire.02|terminate.02|sack.01|can.02 - - -
5 acabar acabar VMN0000 VMN - - 02609764-v - 0 sentence - terminate.01 - - -
6 en en SP SP - - - - 5 cc - - - AM-LOC -
7 el el DA0MS0 DA - - - - 8 spec - - -
8 tracto tracto NCMS000 NC - - 05510907-n - 6 sn - - -
9 digestivo digestivo AQ0MS00 AQ - - 03060385-a - 8 s.a - - -
10 y y CC CC - - - - 5 coord - - -
11 provocar provocar VMN0000 VMN - - 01752884-v - 5 S - provocar.00 - - -
12 problemas problema NCMP000 NC - - 05686955-n - 11 sn - - -
13 digestivos digestivo AQ0MP00 AQ - - 03060385-a - 12 s.a - - -
14 . . Fp Fp - - - - 5 f - - -

```

Con este tipo de archivo para cada una de las frases, se procedió a analizar el *Corpus COVID-19*, estudiando con detenimiento cada ejemplo para etiquetar apropiadamente las categorías en las tres estructuras. Se debe aclarar que los ejemplos para cada verbo se analizaron todos en conjunto, con la idea de dejar analizado cada verbo con todos los posibles sentidos verbales que se iban presentando. En los siguientes apartados se exponen los resultados obtenidos con estos estudios.

#### 2.1.2.2 Estructura-c

Para la estructura-c, los componentes de cada frase se clasificaron en sintagmas mayores (ver 1.8.4.2): SN para el sintagma nominal, SV para el sintagma verbal de los verbos simples y de los compuestos, SP para el sintagma preposicional, SAdv para el sintagma adverbial, SAdj para el sintagma adjetival, SConj para el sintagma de conjunción y O para las oraciones subordinadas que servían como complemento del verbo. En la Tabla 9 se muestra el esquema que se siguió para clasificar cada sintagma que se encontró en las frases. Se optó por la división de SN<sub>n</sub> (SN, SN1, SN2, SN3), porque permitió clasificar todos los posibles sintagmas nominales que se fueron encontrando en las frases del *Corpus COVID-19*. Se intentó dejar el SV en el centro del análisis para mantener el orden prototípico de los constituyentes en español en las tres estructuras. Las categorías de la estructura-c quedaron como se aprecia en la Tabla 9:

**Tabla 9***Categorías de la estructura-c*

ID	Referencia	Frase	LEMA	O	SN	SN1	SV	SN2	SN3	SP	SAdv	SAdj	SConj

El proceso de clasificación de los sintagmas no estuvo carente de decisiones importantes. La primera de ellas fue etiquetar únicamente los constituyentes del verbo en cuestión, para incluir las frases subordinadas, como, por ejemplo, en los casos de oraciones completivas en función de objeto directo o de sujeto, y para analizar también el verbo principal de las frases subordinadas cuando esta frase presentara el verbo que se buscaba estudiar. Vázquez et al. (2006) aplicaron este mismo tipo de análisis en su trabajo para anotar los constituyentes relacionados con el predicado verbal. En el ejemplo de la Tabla 10, el verbo que interesaba era INFECTAR y los elementos que se analizaron pertenecían a la frase subordinada que se subraya, mientras que el resto de esta oración quedó sin etiquetarse. Por razones de espacio, cada tabla con ejemplos incluirá únicamente las etiquetas pertinentes para mostrar la situación encontrada.

**Tabla 10***Análisis en la estructura-c para oración subordinada*

ID	Frase	LEMA	SN	SV	SP	SAdv
301	Que los animales de compañía propaguen el nuevo coronavirus ha sido también negado por la institución, que no tiene ninguna prueba de que <u>el nuevo coronavirus infecte a animales de compañía como los perros y los gatos.</u>	INFECTAR	el nuevo coronavirus	infecte	a animales de compañía	como los perros y los gatos

Otra decisión importante fue dejar los verbos auxiliares y los modales como parte del SV, y analizar la semántica del verbo principal y no la del modal o auxiliar. Al SV lo componen entonces todas las formas verbales del verbo que se buscaba estudiar, como en la forma simple de la Tabla 10 y la compuesta de la Tabla 11:

**Tabla 11***Clasificación del sintagma verbal SV en la estructura-c*

ID	Frase	LEMA	SN	SV	SN2	SP	SAdv
828	La pandemia también ha afectado al entorno de trabajo de las mujeres dominicanas.	AFECTAR	la pandemia	ha afectado		al entorno de trabajo de las mujeres dominicanas	también
706	<u>Todas las vacunas pueden causar algunos efectos secundarios [...]</u>	CAUSAR	todas las vacunas	pueden causar	algunos efectos secundarios		

Continuando con la clasificación, se decidió separar el SP, el SAdv y el SConj de las oraciones subordinadas y de infinitivo que complementaban al verbo, con la idea de que la oración se distinguiera de la preposición del SP, del adverbio del SAdv y de la conjunción del SConj, como se aprecia en la Tabla 12:

**Tabla 12**

*Etiquetado de las oraciones subordinadas en la estructura-c*

ID	Frase	LEMA	O	SN	SV	SN2	SP	SAdv	SConj
861	Los estudios muestran que las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19, incluso para personas con alto riesgo de contraer la enfermedad.	<b>MOSTRAR</b>	las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19	los estudios	muestran			incluso para personas con alto riesgo de contraer la enfermedad	<b>que</b>
565	[...] virus de animales que usan ácido ribonucleico (ARN) como material genético para mutar [...]	<b>USAR</b>	mutar	virus de animales	usan	ácido ribonucleico ARN	<b>para</b>	como material genético	
575	El cuidador debe usar una mascarilla facial cuando esté cerca de la persona enferma [...].	<b>USAR</b>	esté cerca	El cuidador	debe usar	una mascarilla facial	de la persona enferma	<b>cuando</b>	

Cabe señalar que, en esta estructura o nivel superficial, la tesiaría no clasificó los adjuntos ni los argumentos, sino que trabajó únicamente con las etiquetas sintácticas para los sintagmas de los constituyentes. Con las categorías para la estructura-c definidas, se procedió a establecer lo que serían las funciones gramaticales que respondieran a la estructura-f y a las necesidades encontradas en el corpus.

### 2.1.2.3 Estructura-f

La estructura-f se trabajó con las categorías que se mencionaron en la metodología y que se muestran aquí en la Tabla 13, en donde se siguen las funciones propuestas en (7) en el marco teórico por LFG, con la salvedad de que se agregaron otras etiquetas, como XCOMP Prep para los suplementos y los argumentos regidos por el verbo para materia o tema, PRED para los predicados y atributos, CONJ para las conjunciones y la conjunción completiva ‘que’ de las oraciones subordinadas, y CAG para el complemento agente en oraciones pasivas.

**Tabla 13***Clasificación de la estructura-f*

ID	Referencia	Frase	LEMA	SUJ	V	OBJ	OBJ0	OBL0	XCOMP OBJ	XCOMP Prep	PRED	ADJUNTO	CONJ	CAG
----	------------	-------	------	-----	---	-----	------	------	-----------	------------	------	---------	------	-----

Martí Antonín et al. (2007) anotaron automáticamente las funciones gramaticales con las etiquetas SUJ, CD, CI, CREG, CAG, CPRED, ATR, CC, CCT y CCL, como se explicó en el estado de la cuestión. Si se comparan las etiquetas de ese estudio con las de esta tesis, se determina que hay similitudes en las categorías etiquetadas, pero con la diferencia de que, en este nivel, Martí Antonín et al. (2007) clasificaron los complementos circunstanciales de tiempo y de lugar, mientras que, en la presente tesis, los complementos circunstanciales se dejaron todos como ADJUNTOS en este nivel para clasificarse posteriormente con etiquetas semánticas en la estructura-a. Marimon (2015), al trabajar con la Gramática Sintagmática Nuclear (*Head-driven phrase structure grammar* o *HPSG*), clasificó los modificadores y los adjuntos en la estructura de las dependencias, como se mostró en la Figura 5 del estado de la cuestión, y en la semántica que se mostró en la Figura 6.

En la Tabla 14 se presenta el etiquetado para la frase con ID 861, para la cual se mostró el etiquetado de su estructura-c en la Tabla 12 del apartado anterior. Esta frase tiene una oración subordinada (O), introducida por la conjunción completiva ‘que’, que actúa como complemento de objeto directo (XCOMP OBJ) en la estructura-f. Como se explicó en el apartado anterior, se decidió separar la conjunción completiva de la frase subordinada, quedando ‘que’ etiquetado como una conjunción completiva y no como parte de la oración subordinada.

**Tabla 14***Etiquetado de XCOMP en la estructura-f*

ID	Frase	LEMA	SUJ	V	XCOMP OBJ	ADJUNTO	CONJ
861	Los estudios muestran <u>que las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19</u> , incluso para personas con alto riesgo de contraer la enfermedad.	MOSTRAR	los estudios	muestran	las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19	incluso para personas con alto riesgo de contraer la enfermedad	que

El XCOMP Prep es una etiqueta para los complementos preposicionales, de régimen preposicional o también llamados suplementos, que ciertos verbos requieren para completar su predicado; estos no son complementos circunstanciales, aunque *FreeLing* los etiquete como tales. Un ejemplo de este fenómeno se encuentra en la Tabla 15:

**Tabla 15**

*Etiquetado de XCOMP Prep en la estructura-f*

ID	Frase	LEMA	SUJ	V	OBJ	XCOMP Prep
2949	<u>Las vacunas protegen a las personas de infectarse [...]</u> .	PROTEGER	Las vacunas	protegen	a las personas	de infectarse

‘CAG’ es una etiqueta ampliamente utilizada para el complemento agente en las oraciones pasivas e indica el agente de la acción como Martí Antonín et al. (2007) emplearon también. Únicamente se encontraron cuatro casos para esta etiqueta en las 801 frases y estos se aprecian en la Tabla 16:

**Tabla 16**

*Casos de ‘CAG’ en el Corpus COVID-19*

ID	Frase	LEMA	SUJ	V	ADJUNTO	CAG
3128	Las máscaras quirúrgicas (normales) son utilizadas desde hace tiempo por personas sanas en Hong Kong, Japón, Singapur y Malasia.	UTILIZAR	Las máscaras quirúrgicas normales	son utilizadas	desde hace tiempo / en Hong Kong, Japón, Singapur y Malasia	por personas sanas
4453	Para febrero de 2021 diez vacunas han sido autorizadas para su uso público por al menos una autoridad reguladora competente.	AUTORIZAR	diez vacunas	han sido autorizadas	Para febrero de 2021 / para su uso público	por al menos una autoridad reguladora competente
4454	Esta terapia fue autorizada por la FDA estadounidense en agosto de 2020 como tratamiento de emergencia contra la COVID-19.	AUTORIZAR	Esta terapia	fue autorizada	en agosto de 2020 / como tratamiento de emergencia contra la COVID-19	por la FDA estadounidense
5533	Pero ahora, la tendencia cambió y <u>la tos ha sido superada por dolor de cabeza (66%), dolor de garganta (53%), secreción nasal (49%) y fiebre (42%) en menores de 40 años</u> , mientras que los mayores de 40 mostraron resultados similares, precisó el profesor Spector.	SUPERAR	la tos	ha sido superada	en menores de 40 años	por dolor de cabeza 66 %, dolor de garganta 53%, secreción nasal 49% y fiebre 42%

Con las etiquetas anteriores se clasificaron las funciones gramaticales encontradas en el *Corpus COVID-19*. Con este análisis se procedió luego a etiquetar los roles temáticos de los argumentos y de los adjuntos encontrados en cada frase.

#### 2.1.2.4 Estructura-a

Definir las etiquetas para los roles temáticos de los argumentos y de los modificadores verbales fue una tarea que requirió tomar en cuenta las necesidades del corpus y el material ya explotado en otros trabajos. La investigadora de esta tesis mantuvo las categorías ordenadas para cada estructura, de manera que no fueran a causar confusión en el etiquetado y para dejar claro cada análisis desarrollado. Para esto, tomó como base los roles temáticos propuestos por LFG en la Tabla 1 y en (10) del marco teórico, y las reglas de jerarquía ahí expuestas para evitar etiquetas ambiguas. En el estudio de Reyes et al. (2013), se combinaron etiquetas de roles temáticos con funciones gramaticales cuando emplearon «Evento:Objeto (EO)» para designar, posiblemente, el rol temático menos agentivo (tema o paciente). Sin embargo, ‘objeto’ no es un rol temático, sino una función gramatical que debe indicarse en tal análisis y no dentro del semántico, aunque se haya empleado para anotar automáticamente varios roles con una misma etiqueta.

Similar a Reyes et al. (2013), en García Marchena (2009) se mezclaron categorías gramaticales («*attribute*») con roles temáticos, porque con el método automatizado que emplearon se etiquetaron los roles en inglés «*agent*», «*theme*» y «*patient*» para los complementos y «*attribute*», «*manner*», «*time*» y «*place*», para los adjuntos. Con la Teoría Sentido-Texto, García Marchena (2009) apuntó que se sacrifica la complejidad semántica en miras a la automatización del etiquetado, cuando se limita la anotación del sujeto a «*agent*» únicamente, dejando de lado otras construcciones en donde el sujeto puede ser <paciente>, <fuente> o <experimentador>, por ejemplo.

Para asegurarse de que en esta tesis no hubiera ambigüedad en sus etiquetas, la investigadora también estudió las guías de anotación de *PropBank* (Bonial et al., 2015). Estas etiquetas se modificaron para que concordaran con las características de la lengua española, dado que fueron creadas para el idioma inglés. El uso de *ProBank* permitió, además, enriquecer en detalle el etiquetado de los roles semánticos, porque ofreció a la investigadora el ambiente preciso para profundizar en los análisis de los argumentos y de los adjuntos verbales. A diferencia de otros estudios, como en García Marchena (2009), Reyes et al. (2013) y López Rodríguez (2020), en esta

tesis no se trabajó con proyecciones semánticas, ni con macro o proto roles. En el apartado de la metodología, en los cuadros 1 y 2, se dieron a conocer los roles temáticos con los que la tesaria desarrollaría el análisis detallado y la anotación del corpus. A continuación, se presentan los roles resultantes con los que se trabajó:

- (a) **ARG0: agente.** Entidad que realiza la acción descrita por el verbo y que presenta los rasgos de [+volitiva], [+animada], [+intencionalmente].
- (b) **ARG01: causa.** Entidad que causa una acción en otras entidades y que presenta los rasgos de [-volitiva], [-animada], [-intencionalmente].
- (c) **ARG1: paciente.** Entidad animada que recibe la acción descrita por el verbo y que presenta los rasgos de [+animado] y, generalmente, se evidencia en un objeto directo.
- (d) **ARG2: tema.** Entidad no animada que recibe la acción descrita por el verbo y que presenta los rasgos de [-animado].
- (e) **ARG3: beneficiario.** Entidad que recibe consecuencias positivas o negativas de la acción descrita por el verbo y que presenta los rasgos de beneficios recibidos y, generalmente, se evidencia en un objeto indirecto.
- (f) **ARG4: instrumento.** Entidad que sirve de instrumento o medio en la realización de una acción y que presenta los rasgos de [-animada], [-volitiva].
- (g) **ARG5: experimentador.** Entidad que percibe o experimenta sentimientos o sensaciones inspirados o provocados por una fuente, esta ha sido afectada por la inspiración y presenta los rasgos de [-rol activo] y, generalmente, se asocia con verbos de percepción y sensación.
- (h) **ARG-COM: comitativo.** Entidad que acompaña a otra en la acción del verbo y que presenta los rasgos de [+animada], [+volitiva].
- (i) **ARG-GOL: meta.** Punto final o lugar de destino hacia donde se produce un movimiento y se refiere a llegadas o a logros.
- (j) **ARG-DIR: direccional.** Entidad que inspira o provoca sentimientos o sensaciones en otra, presenta los rasgos de [+rol activo] en la inspiración y, generalmente, se asocia con verbos de percepción y sensación. Es también el punto inicial del movimiento o de una situación y se presenta como lugares de partidas.
- (k) **ARG-LOC: lugar.** Lugar relacionado con el desarrollo de la acción.

- (l) **ARG-TMP: tiempo.** Punto o período de tiempo relacionado con el desarrollo de la acción, ya sea como principio, período de desarrollo o final y, generalmente, se presenta como fechas.
- (m) **ARG-EXT: extensión.** Se refiere a la cantidad de cambio ocurrida de una acción y se presenta, generalmente, con adjuntos numéricos, cuantificadores y con comparativos.
- (n) **ARG-DIS: discursivo.** Se refiere a los conectores discursivos.
- (o) **ARG-MOD: modal.** Se refiere a los auxiliares del verbo, sean modales o aspectuales.
- (p) **ARG-ADV: adverbial.** Se refiere a los elementos sintácticos que modifican la estructura del evento verbal y que no cabe dentro de las descripciones de las otras etiquetas. Se presentan en modificadores temporales (gerundios), intensionales ('posiblemente', etc.), de enfoque ('únicamente', 'incluso', etc.) u oracionales ('francamente hablando', 'afortunadamente', etc.).
- (q) **ARG-PRP: finalidad.** Se refiere a las cláusulas que muestran la motivación de la acción del verbo.
- (r) **ARG-NEG: negación.** Se refiere a los marcadores adverbiales de negación.
- (s) **ARG-CAU: causativa.** Se refiere a las cláusulas que indican la razón de la acción.
- (t) **ARG-MNR: manera.** Se refiere a los adverbios de manera que especifican cómo una acción se lleva a cabo.
- (u) **ARG-REC:** se refiere a los pronombres de recíproco, reflexivo, pronominal e impersonal; también a los clíticos.

Marimon (2015) (ver la Figura 6 en el estado de la cuestión), distinguió los complementos de los adjuntos. A diferencia del análisis de Marimon (2015) estos dos componentes de la oración se etiquetan en detalle en la presente tesis. Aunque es importante clasificarlos en estas dos categorías, la anotación de textos para ML en AI resulta más ventajosa cuando cuenta con más datos lingüísticos, como los proporcionados con los análisis de esta tesis. Con las definiciones de las etiquetas en (a) - (u) se dio lugar a la anotación de los argumentos y de los modificadores verbales en la estructura-a. En adelante, se usará la etiqueta asignada al rol entre <...> (Bresnan et al. (2016) o el nombre del rol en sí intercambiamente, dependiendo de cómo presente más claramente la explicación.

En el apartado de la metodología se detalló que se emplearían como guía las anotaciones automáticas de la herramienta *FreeLing* o *TextServer*. Sus etiquetas no son perfectas y tienden a

presentar muchos errores, principalmente para los análisis en español, como se ha comprobado también en otros trabajos (Castellón et al., 2012); por esta razón, la tesiaría revisó detenidamente cada una de las que el *parser* asignó para asegurarse de que cuando se fijaran las funciones gramaticales y luego los roles temáticos en el corpus, estos fueran realmente los adecuados para cada caso. Algunos ejemplos de etiquetados erróneos por parte de *TextServer/FreeLing* se muestran en (22)-(27), junto con la decisión que la tesiaría tomó para solucionar el error:

- (22) Para la frase con ID 1793 «los factores inflamatorios asociados con las enfermedades que contienen principalmente IL-6 aumentaron significativamente», *TextServer* etiquetó el argumento “los factores inflamatorios asociados con las enfermedades que contienen principalmente IL-6” como <A0> en lugar de <A1>. La tesiaría lo etiquetó como <tema>.
- (23) Para la frase con ID 1875 «Los expertos no ven necesaria la cuarta dosis para todos», *TextServer* etiquetó el argumento “Los expertos” como <A0>, cuando en el sentido que tiene la frase el sujeto es más un <experimentador> que un <agente>. La tesiaría lo etiquetó como <experimentador>, tomando en cuenta la jerarquía de prominencia para los roles temáticos que proponen Bresnan et al. (2016) en (10).
- (24) Para la frase con ID 3158 «La India registra casi 200.000 nuevos casos de coronavirus en las últimas 24 horas», *TextServer* etiquetó el argumento “en las últimas horas” como <AM-LOC> y “casos de coronavirus” como <AM-TMP>. La tesiaría los etiquetó como <tiempo> y <tema>, respectivamente.
- (25) Para la frase con ID 4143 «Sorprendentemente, las complicaciones tromboticas apenas se han descrito», *TextServer* etiquetó “Sorprendentemente” como <AM-TMP>, cuando este no es un adverbio de tiempo. La tesiaría lo etiquetó como el adjunto <ARG-ADV>.
- (26) Para la frase con ID 4568 «El gobierno de Hong Kong advirtió a cualquiera que viaje fuera de la ciudad que no toque animales», *TextServer* etiquetó ambos argumentos “a cualquiera que viaje fuera de la ciudad” y “que no toque animales” como <A1>, cuando “a cualquiera que viaje fuera de la ciudad” es un objeto indirecto, por lo que no puede ser <paciente> o <tema>. La tesiaría etiquetó como <tema> al OBJ y <beneficiario> al OBJ $\theta$ .

- (27) Para la frase con ID 5457 «Debido a la complejidad y coste de las pruebas PCR, en los primeros meses de la pandemia casi ningún país dispuso de datos confiables sobre la prevalencia del virus en su población», *TextServer* etiquetó “en los primeros meses de la pandemia” como <AM-LOC>, cuando se trata de <ARG-TPM>. La tesiaría lo etiquetó como <ARG-TPM>.

Este tipo de errores se repitieron en otras frases que, por espacio, no se expondrán aquí. Pero, lo que resulta imprescindible subrayar es que, por este tipo de discrepancias, los resultados dados por este *parser* no se tomaron como ciertos hasta haberse comprobado de que lo fueran. En varias frases se tuvo que reconstruir la estructura de los constituyentes original, porque el *TextServer* no comprendía el orden de los argumentos y los etiquetaba erradamente, como en (28):

- (28) Para la frase con ID 5628 «Esto se debe sobre todo a la dependencia económica y a las dificultades para denunciar que han provocado e incrementado las nuevas medidas de confinamiento y aislamiento social», *TextServer* etiquetó “la dependencia económica y a las dificultades para denunciar” como <A1> y “las nuevas medidas de confinamiento y aislamiento social” también como <A1>, cuando este último argumento debe ser <A0>. La tesiaría lo etiquetó como <causa>.

En casos como el de (23), la tesiaría varias veces tuvo que cavilar por horas y revisar otros estudios para conseguir los roles temáticos que fueran realmente pertinentes al sentido y la definición del verbo en la frase estudiada. Uno de estos otros casos fue con ‘contraer una enfermedad’, que *TextServer* etiqueta al sujeto como <A0>, mientras que *GRIAL* lo etiqueta como <iniciador>. La investigadora considera que ninguno de los roles parece ser pertinente a la acción en donde un sujeto adquiere una enfermedad, como la semántica que los siguientes ejemplos exponen (29-32):

- (29) ID 3557 «Las personas que contraen el coronavirus también pueden tener diarrea [...]».
- (30) ID 3560 «A veces, las personas que están completamente vacunadas contraen una infección irruptiva [...]».
- (31) ID 3570 «Cualquier persona puede contraer la COVID-19 [...]».
- (32) ID 3576 «Asimismo, se ha confirmado que otros animales –como perros, gatos, tigres, leones y murciélagos– pueden contraer el COVID-19 al igual que los seres humanos».

De manera que, basándose en la jerarquía de prominencia de los roles temáticos se decidió dejar la etiqueta <beneficiario> para los sujetos de estas cuatro frases, porque no hay volición en el sujeto que adquiere la enfermedad para que este sea <agente>, y tampoco es el origen (<iniciador>) de la enfermedad, sino que la enfermedad le llega al sujeto.

Otro caso fue el del verbo RESPONDER cuando está en forma intransitiva y adquiere un sentido más cercano al de CORRESPONDER, como en la frase con ID 4889:

- (33) «Una enfermedad, incluida la COVID-19, no siempre responde a los mismos síntomas». El rol temático del sujeto es <tema> (Drăgușin, 2015) y no <agente>, como lo etiquetó el *parser* automático.

Las frases de DURAR fueron otro ejemplo de decisión entre lo que anotó *TextServer* y *GRIAL*, dado que el primero etiquetó al sujeto de las 4 oraciones como <A0> o <A1>, y *GRIAL* como <iniciador>. Para su adjunto, *TextServer* lo anotó como <A1>, <A4>, <AM-TMP> y <A2>, mientras que *GRIAL* lo deja como <cantidad>. La tesiaría decide dejar el par de roles temáticos para este verbo como <ARG-DIR, ARG-EXT>, en donde el primero es para el sujeto y el segundo para su adjunto.

VISITAR también requirió un análisis profundo. Para la frase con ID 4780 (34), *FreeLing* etiquetó el OBJ como <A1>, mientras que la tesiaría considera que este argumento cumple más la función semántica de <ARG-LOC>:

- (34) «Use una mascarilla de tela con al menos 2 capas cuando visite el consultorio o la sala de emergencia [...]».

Joan Casademont (2020) concuerda con esta etiqueta cuando afirma que para este verbo ella encontró que los roles temáticos más pertinentes para el objeto directo de este verbo eran <paciente-tema/lugar>, dependiendo del SN que tenga ese argumento. De manera que para algunos casos el OBJ será <tema> como en las frases con ID 4752 (35), 4787 (36) y 4790 (37), o <lugar> como en la frase con ID 4780 (34), mencionada anteriormente.

- (35) «Ir al médico en caso de fiebre, tos y dificultad para respirar, llamando con antelación si se encuentra en zonas donde se está propagando el virus o si se las han visitado en los últimos 14 días».
- (36) «Informó de que había visitado regularmente un mercado local de productos frescos en Wuhan antes de la aparición de la enfermedad el 5 de enero de 2020».

- (37) «La bestia sale de noche, dijo Cuomo, refiriéndose a los escalofríos, dolores en el cuerpo y fiebre alta que lo visitaron el 1 de abril».

A modo de resumen cuantitativo, en las Tablas 17 y 18 se muestra la distribución de los ejemplos para cada uno de los roles temáticos que fueron anotados por la tesaria en el análisis de la estructura-a. Este conteo no se consideró necesario para las estructuras-c y -f, dado que la meta final de la anotación de esta investigación son los roles temáticos. La Tabla 17 presenta los resultados para los argumentos primarios (Bonial et al., 2015), mientras que en la Tabla 18 se dan los de los modificadores del predicado verbal. Para los argumentos primarios se puede apreciar que el rol con más ejemplos fue <tema>, mientras que <instrumento> fue el menos encontrado. Asimismo, por la temática del virus del SARS-CoV-2, se hallaron múltiples ejemplos para los adjuntos de tiempo y de lugar, con 179 y 104, respectivamente para cada uno (Tabla 18).

**Tabla 17**

*Distribución de etiquetas para los argumentos primarios*

ARG0 agente	ARG01 causa	ARG1 paciente	ARG2 tema	ARG3 beneficiario	ARG4 instrumento	ARG5 experimentador
133	163	45	651	25	14	59

**Tabla 18**

*Distribución de etiquetas para los modificadores verbales*

ARG TMP	ARG LOC	ARG COM	ARG EXT	ARG DIR	ARG GOL	ARG REC	ARG PRP	ARG CAU	ARG DIS	ARG MOD	ARG NEG	ARG ADV	ARG MNR
179	104	11	47	46	58	143	52	31	35	216	44	136	85

En los siguientes apartados se darán a conocer los resultados de la creación y aplicación del modelo de anotación para los roles temáticos del *Corpus COVID-19* y, posteriormente, los resultados del acuerdo entre los anotadores.

### **2.1.3 Desarrollar un modelo para el etiquetado de roles temáticos**

En el apartado 1.8.4.3 de la metodología se explicaron los pasos que se siguieron para crear el modelo y la guía de anotación que se entregó a los anotadores. De manera que lo primero que se hizo fue fusionar los roles temáticos de la estructura-a para los argumentos verbales primarios y dejar únicamente los adjuntos de lugar y de tiempo:

- (a) **Agente:** entidad que realiza la acción descrita por el verbo o causa una acción en otras entidades. Fusiona los roles de agente y causa.

- (b) **Tema:** entidad que recibe la acción descrita por el verbo. Fusiona los roles de tema y paciente.
- (c) **Meta:** El punto final o lugar hacia donde se dirige una acción descrita por el verbo con consecuencias positivas o negativas. Fusiona los roles de beneficiario y meta, en donde se consideran las llegadas, los logros y los beneficios.
- (d) **Experimentador:** entidad que recibe o experimenta sentimientos o sensaciones inspirados o provocados por una fuente. Posee un rol pasivo en los verbos de percepción y sensación.
- (e) **Instrumento:** entidad que sirve de instrumento o compañía a otra en la realización de una acción.
- (f) **Fuente:** punto inicial del movimiento o de una situación, también entidad que inspira o provoca sentimientos o sensaciones en otros. Fusiona los roles de fuente y origen. Posee un rol activo en la inspiración, como en verbos de percepción y sensación.
- (g) **Lugar:** es donde se desarrolla la acción.
- (h) **Tiempo:** punto o período en el tiempo relacionado con el desarrollo de la acción, ya sea como principio, origen, periodo en desarrollo o final. Son fechas o expresiones de tiempo.

La distribución de los ejemplos anotados por la tesiaría para cada uno de estos roles fusionados se presenta en la Tabla 19, manteniendo <tema> como el rol con más ejemplos anotados e <instrumento> como el menos etiquetado:

**Tabla 19**

*Distribución de las etiquetas por rol temático*

agente	tema	meta	experimentador	fuentes	instrumento	lugar	tiempo
295	756	82	59	35	25	103	176

El proceso de desarrollo del modelo de anotación no estuvo exento de dificultades que se tuvieron que resolver prontamente. El primer obstáculo que se tuvo que superar fue haber encontrado un error manual por parte de la tesiaría cuando fijó el número de ID a cada frase del *Corpus COVID-19*. Este error causó que el mismo ID se encontrara en dos frases diferentes. Este problema se tuvo que corregir manualmente también, porque cuando se intentó automatizar la corrección para asignar el número correcto a las frases con el problema, Excel aplicaba un tipo de

aleatoriedad, causando que, aunque se ordenara por verbo, las frases quedaran en un orden distinto cada vez que se ordenaban. Este proceso costó varias horas de trabajo a la tesiaría.

Otra situación que consumió tiempo fue corregir la inconsistencia en Excel para mantener el formato dado a las frases, principalmente, cuando se había subrayado el argumento por analizar en cada una de las 1531 oraciones del modelo. Si se aplicaba “copiar” y “pegar” el texto en otra columna, unas veces llevaba consigo el formato, mientras que otras veces no, especialmente si la frase era extensa. Se probó el copiar las tablas por medio de varias opciones del programa de Excel, pero era impredecible el resultado cada vez que se debía crear otra hoja de Excel porque se repetía el problema. Las listas de la encuesta se tuvieron que enmendar múltiples veces, porque, en algunos casos, los anotadores encontraron frases que no tenían el argumento subrayado y, a causa de esto, no podían responder esa frase en la encuesta.

Excel también causó otro problema bastante serio cuando se estaba anotando el corpus y cuando se estaba corrigiendo el error con el ID para las frases. En al menos tres ocasiones, el programa se reinició a la versión anterior de la hoja de trabajo, borrando horas de análisis del documento que se tenía abierto. A pesar de que la tesiaría guardaba copias con versiones diarias del trabajo, una vez que esto ocurría, no se podía recuperar el avance de ese día, a pesar de autoguardarse el documento. El problema se pudo solucionar con un cambio de computadora, aunque este cambio se dio por otras razones técnicas. Al parecer, este problema de reinicio de Excel no es insólito, según se consultó en foros de usuarios del programa. Por las razones anteriores, la tesiaría considera que Excel no fue la mejor herramienta para haber trabajado con datos lingüísticos, particularmente cuando se debía mantener algún formato de escritura.

En el apartado de la metodología se explicó que para que cada frase fuera evaluada por al menos tres anotadores distintos, la lista de 1531 argumentos se multiplicó por tres y el total de 4593 ítems se dividió entre los 15 anotadores humanos para que cada persona anotara  $\approx 306$  argumentos y para que ningún anotador etiquetara el mismo argumento dos veces. Sí sería probable que los anotadores evaluaran la misma frase, pero no para etiquetar el mismo argumento. En la Tabla 20 se aprecia esta distribución:

**Tabla 20**

*Distribución de los ítems según el anotador*

Anotador	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Número de ítems	305	306	306	306	306	305	306	306	306	306	307	307	307	307	307

Los 15 anotadores se dividieron a su vez en cinco grupos, con tres integrantes cada uno, los cuales quedaron repartidos como se muestra en la Tabla 21. De los 15 estudiantes del posgrado que se contaban como posibles anotadores, dos de ellos no se reportaron cuando se asignó el número de anotador y su correspondiente encuesta. Por esta razón, la tesiaría encontró otros dos anotadores, también con formación en lingüística, que aceptaron participar en la encuesta (6e y 15k). A todos los anotadores se les dio aproximadamente una semana para que completaran y entregaran la encuesta; pero, al acercarse la fecha límite para entregar las encuestas, uno de los dos estudiantes del posgrado que no se habían reportado hasta el momento, solicitó ser parte de la encuesta y el profesor Leoni de León le entregó la carpeta del anotador 15. Cuando todos los anotadores devolvieron los resultados, la tesiaría decidió dejar la carpeta para el anotador 15 con dos anotadores en lugar de uno solo. Dadas las circunstancias, el grupo 5 quedó con cuatro anotadores para una parte del corpus.

**Tabla 21**

*Distribución de los grupos de anotadores*

Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
anotador 1	anotador 2	anotador 3	anotador 4	anotador 5
anotador 6e	anotador 7	anotador 8	anotador 9	anotador 10
anotador 11	anotador 12	anotador 13	anotador 14	anotador 15 anotador 15k
<b>917 frases</b>	<b>919 frases</b>	<b>919 frases</b>	<b>919 frases</b>	<b>1226 frases</b>

Al analizar las respuestas de los 16 anotadores (ver la Tabla 22), se encontró que el grupo 2 tuvo la mayor cantidad de frases sin respuesta (7 frases), seguido por el grupo 4. En cuanto a las frases con más de una respuesta, el grupo 4 fue el que tuvo más (4 frases). En total, de las 4900 frases analizadas por los anotadores humanos, a 23 de ellas no le dieron respuesta y 9 tuvieron más de una, provocando que estas 32 frases se consideraran nulas.

**Tabla 22**

*Respuestas de los anotadores humanos*

Frases	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Total
Asignadas	917	919	919	919	1226	4900
Sin respuesta	5	7	2	6	3	23
Con más de 1 respuesta	2	0	0	4	3	9

Las respuestas nulas se descartaron y con el resto de las frases se trabajó para calcular el acuerdo entre los anotadores y el desempate para los casos de frases con respuestas en desacuerdo. Estos dos procesos se explicarán el siguiente apartado.

Cuando la tesiera tuvo la oportunidad de entrevistar a *ChatGPT*, esta debió cambiar el formato de subrayado en los argumentos y dejarlos entre asteriscos, como se explicó en [1.8.4.3](#). Este fue otro paso que tomó un tiempo considerable, dado que la encuesta estaba en Excel y, como ya se indicó anteriormente, esta herramienta no se presta para ediciones de textos a gran escala. En consecuencia, una a una, cada frase se cambió para que *ChatGPT* comprendiera la tarea demandada, como se expuso en (17) y (18) del apartado de la metodología.

El proceso de entrevista a *ChatGPT* tuvo su proceso de aprendizaje para poder encontrar el *prompt* que le explicara al LLM la tarea lo suficiente como para que respondiera apropiadamente. En el apartado de la metodología se explicó que no se buscaba entrenar a *ChatGPT* con métodos de *fine-tuning*, sino que lo que se quería era obtener las respuestas del modelo aplicando la misma guía que se había dado a los anotadores humanos con *zero-shot learning*. En ambos casos, ni a los anotadores humanos ni al LLM, se les ofreció explicaciones profundas teóricas con ejemplos de los roles temáticos que debían emplear en la encuesta. Esta parte de la investigación se enfocaba en la intuición para decidir cuál rol era más pertinente en cada caso, dadas las ocho posibles opciones a escoger.

Después de varios intentos, el *prompt* que mejores resultados ofreció fue: «Quiero saber cuál es el rol temático del segmento entre asteriscos de las frases que te voy a ir dando. Para esto quiero que uses únicamente el conjunto de roles temáticos: agente, quien realiza la acción; tema, a quien le ocurre la acción; meta, el lugar o persona hacia donde se dirige un movimiento o acción; experimentador, quien percibe o experimenta sensaciones provocadas por una fuente; instrumento, con qué o con quién se realiza una acción; fuente, provoca sensaciones en otra entidad o es el origen de donde parte una acción; lugar, donde se desarrolla la acción; y tiempo, el período relacionado con el desarrollo de la acción».

En la Tabla 23 se muestran los datos cuantificados de las respuestas de este modelo al *prompt*. Para un total de 1531 argumentos, *ChatGPT* dejó 63 respuestas nulas, un número bajo, considerando la cantidad de ítemes preguntados. Es preciso señalar que cada respuesta, válida o nula, fue documentada manualmente por la investigadora y que esto podría haber dado lugar a errores que no necesariamente fueran causados por el desempeño del modelo. Aquí también jugó

la subjetividad de la tesiaría al tener que decidir si hubo respuesta o no para los casos que quedaron nulos. Se procedió a dejar como válidas las respuestas que mencionaban un rol temático, pero únicamente si era alguno de los ocho establecidos. No obstante, sí se encontraron irregularidades en el modelo, las cuales se detallarán a continuación.

### Tabla 23

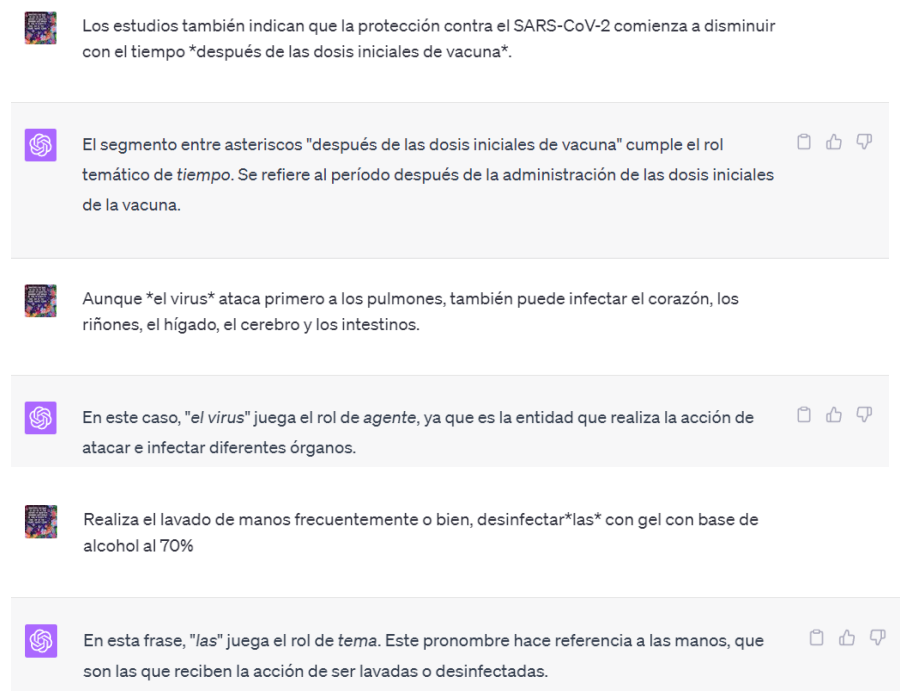
#### Respuestas de ChatGPT a la encuesta

ChatGPT	Número de frases
Frases asignadas	1531
Frases sin respuesta	45
Frases con más de 1 respuesta	18
<b>Total de las respuestas nulas</b>	<b>63</b>

En esta investigación se trabajó con los modelos de *ChatGPT-3.5* y *ChatGPT-4*, que permite 25 interacciones cada tres horas. Cuando se entrevistó a *ChatGPT-4*, las respuestas que dio el modelo mantenían el mismo formato, corto y preciso, con la explicación de por qué elegía el rol temático para ese argumento, como se aprecia en la Figura 28:

### Figura 28

#### Ejemplo de respuesta de ChatGPT-4 para la encuesta



Los estudios también indican que la protección contra el SARS-CoV-2 comienza a disminuir con el tiempo \*después de las dosis iniciales de vacuna\*.

El segmento entre asteriscos "después de las dosis iniciales de vacuna" cumple el rol temático de *tiempo*. Se refiere al período después de la administración de las dosis iniciales de la vacuna.

Aunque \*el virus\* ataca primero a los pulmones, también puede infectar el corazón, los riñones, el hígado, el cerebro y los intestinos.

En este caso, "el virus" juega el rol de *agente*, ya que es la entidad que realiza la acción de atacar e infectar diferentes órganos.

Realiza el lavado de manos frecuentemente o bien, desinfectar\*las\* con gel con base de alcohol al 70%

En esta frase, "las" juega el rol de *tema*. Este pronombre hace referencia a las manos, que son las que reciben la acción de ser lavadas o desinfectadas.

En cambio, cuando se trabajó con el modelo de *ChatGPT-3.5*, que fue la mayoría del tiempo, las respuestas fueron más inconsistentes en cuanto a su formato, a las explicaciones dadas y a la extensión de estas (ver la Figura 29). A este modelo se le debía repetir el *prompt*, más o menos, cada diez interacciones y, de no ser así, el modelo comenzaba a responder con alucinaciones: en algunos casos confirmando el enunciado, corrigiendo la información de la frase o dando datos que no se le habían preguntado (ver la Figura 30). No se puede saber con certeza en este momento si con el modelo *GPT-4* ocurriría lo mismo después de las 25 preguntas que permite cada tres horas. Sin embargo, con las pocas respuestas obtenidas de *GPT-4*, este modelo demostró un comportamiento más cercano al humano en cuanto a la elección de los roles temáticos para cada argumento preguntado. No se mantuvo un inventario de cuántas respuestas se tomaron de cada una de las versiones de *ChatGPT*, lo cual habría sido interesante contrastar después.

## Figura 29

### *Variación de las respuestas de ChatGPT-3.5 con el mismo prompt*

#### **User**

No obstante, \*este tiempo de permanencia\* depende de factores como la temperatura y el tipo de superficie.

#### **ChatGPT**

El segmento entre asteriscos "este tiempo de permanencia" funciona como tema en la oración y se refiere al tiempo que el SARS-CoV-2 puede permanecer en una superficie contaminada.

#### **User**

Si una persona obtiene un resultado positivo en COVID-19, debe llamar al teléfono 900 102 112 para que le recojan \*sus datos\* y le informen de las medidas que debe adoptar.

#### **ChatGPT**

En la frase dada, el segmento entre asteriscos "sus datos" funciona como objeto directo y representa la información personal de la persona que obtuvo el resultado positivo en COVID-19. Por lo tanto, su rol temático es tema.

#### **User**

La pandemia también ha afectado \*al entorno de trabajo de las mujeres dominicanas\*.

#### **ChatGPT**

Rol temático: tema (a quien le ocurre la acción).

### Figura 30

#### *Alucinaciones de ChatGPT*

User

\*El Oso y Madroño de la artista Noelia Jiménez, teñido de una amplia paleta de colores vivos\*,  
representa cómo los trabajadores esenciales han unido a la sociedad española durante la pandemia

ChatGPT

El segmento entre asteriscos "El Oso y Madroño de la artista Noelia Jiménez, teñido de una amplia paleta de colores vivos" funciona como una descripción del sujeto de la oración, "representa cómo los trabajadores esenciales han unido a la sociedad española durante la pandemia". En otras palabras, describe las características del objeto que se está utilizando como símbolo de la unión de la sociedad durante la pandemia.

En otros casos, *ChatGPT-3.5* solicitaba el contexto de la frase para poder asignar el rol temático del argumento, como en los casos de (38) a (40). Sin embargo, no se contaba con este tipo de información y los anotadores humanos tampoco tuvieron acceso a ella. Fue interesante notar que, cuando se le preguntó al modelo *GPT-4* las mismas frases de (38) a (40) en otra ocasión, este respondió sin problemas con un rol temático, fuera su respuesta correcta o no. *ChatGPT* tuvo también problemas para asignar un rol a los argumentos de los verbos que eran parte de una oración subordinada, como en (40).

- (38) «Ir al médico en caso de fiebre, tos y dificultad para respirar, llamando con antelación si se encuentra en zonas donde se está propagando el virus o si se \*las\* han visitado en los últimos 14 días».
- (39) «El hospital informó \*esto\* al CCDC, cuyo equipo en la ciudad inició una investigación».
- (40) «En adición, en la parte de ¿Quiénes se han vacunado? se puede seleccionar el rango de tiempo para el que se quieren \*estos datos\*».
- (41) «Los médicos pueden enviar a los pacientes leves a recuperarse \*en su casa\*, bajo las siguientes condiciones».
- (42) «\*Esto causa irritación y por ello empezamos a toser\*, señala Schaffner».

Al igual que con los anotadores humanos, *ChatGPT* no dio respuestas en algunas frases, o dio dos respuestas distintas para un mismo argumento, como «en su casa» para la frase en (41) o dio tres roles diferentes para la frase en (42). En la Figura 32 se puede notar que para el argumento de (41), *GPT-3.5* respondió que era ambos <meta> y <lugar>, sin definir ni explicar la elección de

uno o el otro. Aquí el modelo empleó otro formato con viñetas para enumerar todos los roles que encontró en esta frase. En (42), la investigadora preguntó tres veces a *ChatGPT* con el mismo *prompt*, para saber si su primera respuesta era fiable y este respondió primero que el argumento señalado era <agente>, luego que era <experimentador> y, finalmente, que era <fuente>. En casos como este, la tesaria documentó todas las respuestas dadas, sabiendo que posteriormente sería considerada una respuesta nula, pero que serviría de ejemplo para demostrar el comportamiento errático del LLM. Como se mencionó párrafos arriba, se consideraron preguntas sin respuesta aquellas en las que el modelo no dio un rol temático para el argumento señalado, aunque haya respondido con alguna alucinación o pidiera más contexto.

### Figura 32

*Dos respuestas de ChatGPT para el mismo argumento “en su casa”*

13. Roles temáticos:

- Agente: Los médicos.
- Tema: Los pacientes leves.
- Meta: Su casa.
- Experimentador: No aparece.
- Instrumento: No aparece.
- Fuente: No aparece.
- Lugar: Su casa.
- Tiempo: No aparece.

Este comportamiento endeble concuerda con lo reportado por Reiss (2023), cuando afirmó con su estudio que las mismas entradas o *prompts* dados a *ChatGPT* ofrecían salidas distintas, particularmente con el modelo *GPT-3.5* y que la constancia de las respuestas del LLM era poco fiable. Para comprobar el trabajo de Reiss (2023), la investigadora de esta tesis probó repitiendo el *prompt* a *ChatGPT*, principalmente cuando la respuesta que este modelo daba no coincidía con lo preguntado para el argumento en cuestión. De manera que, en varios casos, las respuestas del modelo no fueron las mismas ni concordaban entre sí. Para estos casos, se dejó la primera respuesta dada por el modelo, si este era un rol temático para el argumento, o se dejó la segunda o tercera respuesta, si las primeras habían sido una alucinación. No se repitió el *prompt* en todos los casos que no produjo una respuesta válida, porque esto habría extendido el tiempo y los recursos disponibles para esta tarea. En adición a lo anterior, no resulta rentable repetir las mismas interrogaciones para comprobar la validez de lo respondido cuando se lleva a cabo una investigación con 1531 preguntas.

En el siguiente apartado se explicarán los resultados obtenidos para el coeficiente de kappa de Fleiss, con el cual se calculó el acuerdo entre los anotadores y con el que se evaluó el comportamiento de los anotadores humanos y el de *ChatGPT*.

#### **2.1.4 Evaluación del acuerdo entre los anotadores**

El acuerdo entre los anotadores se computó por medio del coeficiente kappa de Fleiss, como se explicó en la metodología. Cabe recordar que este coeficiente se debe calcular por grupo y por número de anotadores, de manera que se pueda obtener el valor real del acuerdo por frase anotada por dos o más anotadores diferentes. Cuando se eliminaron las respuestas nulas, quedaron frases analizadas por un solo anotador, por dos anotadores, por tres anotadores y por cuatro anotadores, para el caso del grupo 5. En la Tabla 24 se dan los resultados de  $\kappa$ , de donde se puede extraer que para el grupo 1, con 3 anotadores asignados, se consiguió un acuerdo de 0.455 entre 2 anotadores para 8 frases etiquetadas, y 0.468 entre 3 anotadores para 296 frases etiquetadas. Con estos valores, se considera que el grupo 1, en promedio, tuvo un valor de  $\kappa$  de 0.462. De acuerdo con la Tabla 4, este valor de  $\kappa$  tiene un nivel moderado de acuerdo.

**Tabla 24**

*Kappa de Fleiss para los anotadores humanos*

<b>Grupo</b>	<b>Número de anotadores</b>	<b>Número de frases</b>	<b>Kappa de Fleiss</b>
1	2	8	0.454545455
	3	296	0.468103397
2	2	10	0.367088608
	3	300	0.302744709
3	2	4	0.076923077
	3	303	0.52667051
4	2	9	0.073529412

	3	297	0.588488049
5	3	9	0.524647887
	4	299	0.407739077
<b>Promedio</b>			0.342356638

Para el grupo 2, sin embargo, los valores de  $\kappa$  son los más bajos, con 0.367 entre 2 anotadores para 10 frases, y 0.303 entre 3 anotadores para 300 frases, para un promedio de 0.335 que caería en el nivel justo. El grupo 3 obtuvo 0.077 entre 2 anotadores para 4 frases, para un nivel ligero, mientras que consiguió un valor para  $\kappa$  de 0.527 entre 3 anotadores para 303 frases, el cual está dentro del nivel moderado. El grupo 3 quedó con un promedio de 0.302 para un nivel justo. El grupo 4 presentó un comportamiento similar al del grupo 3, en donde obtuvo un valor de  $\kappa$  de 0.074 entre 2 anotadores para 9 frases, y 0.588 entre 3 anotadores para 297 frases, para dejar un promedio de 0.331. En cuanto al grupo 5, los valores de  $\kappa$  que logró fueron los más altos, con 0.525 entre 3 anotadores para 9 frases, y 0.408 entre 4 anotadores para 299 frases, con lo que consigue un promedio de 0.466 o un nivel moderado. En resumen, el promedio del acuerdo entre los anotadores humanos con formación en lingüística es de 0.342, lo cual lo deja en el nivel justo, de acuerdo con valores de Landis y Koch (1977) de la Tabla 4.

Cuando se combinan las respuestas de *ChatGPT* con las dadas por los humanos, se determina que, en general, los valores de  $\kappa$  tienden a reducirse, como se aprecian en la Tabla 25. A continuación se analizarán en detalle los valores obtenidos.

### Tabla 25

*Kappa de Fleiss para anotadores humanos junto con ChatGPT*

Grupo	Número de anotadores	Número de frases	Kappa de Fleiss
1	2	6	-0.111111111
	3	18	0.464655172
	4	286	0.441196967

2	2	2	1.000000000
	3	26	0.246308113
	4	281	0.357644033
3	2	1	-1.000000000
	3	15	0.312573443
	4	292	0.485387507
4	2	3	0.076923077
	3	23	0.355494505
	4	284	0.521588199
5	2	2	0.200000000
	3	3	0.196428571
	4	22	0.243142724
	5	283	0.380558015
<b>Promedio</b>			0.260674326

Cuando se calculó el coeficiente de kappa de Fleiss para el grupo de humanos, se descartaron también frases que reportaron la respuesta de un solo anotador, porque no había forma de comparar un acuerdo. Sin embargo, cuando se combinaron las respuestas humanas con las de *ChatGPT*, se pudo evaluar las etiquetas dadas para esas frases originalmente descartadas en cada uno de los grupos. El acuerdo registrado para estos casos con ahora 2 anotadores fue de nivel ligero con pocas frases en común para comparar en cada grupo: se alcanzó un valor en promedio de  $\kappa$  de 0.03316239 (-0.111 para el grupo 1; 1 para el grupo 2; -1 para el grupo 3; 0.077 para el grupo 4; y 0.200 para el grupo 5), como se aprecia en la Tabla 26:

**Tabla 26***Acuerdo entre 1 anotador humano y ChatGPT*

Grupo	1 anotador humano + GPT	Número de frases anotadas
1	-0.111	6
2	1	2
3	-1	1
4	0.077	3
5	0.200	2

Cuando se comparan las respuestas de *ChatGPT* con 2, 3 y 4 anotadores en cada grupo, los resultados resultan ser más sugestivos, como se demostrará a continuación. En la Tabla 27 se aprecian los cambios positivos en el valor de  $\kappa$  cuando se calculó el acuerdo entre 2 anotadores humanos y *ChatGPT*, particularmente para los grupos 4 y 5, que pasaron de tener un coeficiente de kappa de Fleiss de 0.077 y 0.074 a 0.313 y 0.355, respectivamente. Claramente esto representa una mejoría en el acuerdo entre los anotadores al contar con un participante más. El grupo 2, sin embargo, bajó el valor de  $\kappa$  de 0.367 a 0.246, mientras que el grupo 1 se mantuvo casi sin variaciones, pero con un leve aumento de 0.455 a 0.465.

**Tabla 27***Comparación del valor de  $\kappa$  al combinar ChatGPT con 2 anotadores humanos*

Grupo	2 anotadores humanos	2 anotadores humanos + GPT
1	0.454545455	0.464655172
2	0.367088608	0.246308113
3	0.076923077	0.312573443
4	0.073529412	0.355494505
5	(no hubo casos)	0.200000000

El acuerdo entre 3 anotadores humanos y *ChatGPT* también mostró cambios, pero estos fueron más bien negativos (ver la Tabla 28), con excepción del grupo 2, que aumentó el valor de  $\kappa$  de 0.303 a 0.358. El grupo 1 redujo su acuerdo de 0.468 a 0.441; el grupo 3, de 0.527 a 0.485; el grupo 4, de 0.588 a 0.522; y el grupo 5, de 0.525 a 0.243, el cual fue el más afectado con la participación de *ChatGPT*:

**Tabla 28**

*Comparación del valor de  $\kappa$  al combinar ChatGPT con 3 anotadores humanos*

Grupo	3 anotadores humanos	3 anotadores humanos + GPT
1	0.468103397	0.441196967
2	0.302744709	0.357644033
3	0.52667051	0.485387507
4	0.588488049	0.521588199
5	0.524647887	0.243142724

Finalmente, cuando se compara el desempeño de *ChatGPT* junto con 4 anotadores humanos del grupo 5, se obtiene también una reducción en el valor del acuerdo, que bajó de 0.408 a 0.381, en conformidad con los datos de la Tabla 29:

**Tabla 29**

*Comparación del valor de  $\kappa$  al combinar ChatGPT con 4 anotadores humanos*

Grupo	4 anotadores humanos	4 anotadores humanos + GPT
5	0.407739077	0.380558015

De acuerdo con los datos anteriores, se considera que el desempeño de *ChatGPT*, en general, fue inferior al de los anotadores humanos, al menos desde la perspectiva del acuerdo entre las respuestas dadas a la encuesta de roles temáticos para el *Corpus COVID-19*. Para respaldar esta afirmación, la investigadora calculó el acuerdo entre las anotaciones teóricas manuales que ella llevó a cabo, las dadas por los anotadores humanos y las dadas por el LLM. Para comparar las anotaciones de los humanos con los de la tesiaría, se desempataron las respuestas que no obtuvieron acuerdo, como se explicó en la metodología, de manera que se estableció una única respuesta para cada argumento de la encuesta. En las frases en donde hubo una mayoría para una etiqueta, se dejó esa etiqueta con mayoría como respuesta. Cuando no hubo una mayoría de votos, el anotador desempataador asignó una, basando su respuesta en las etiquetas asignadas por los otros anotadores. La lista final con las etiquetas asignadas de los anotadores humanos para los 1531 argumentos obtuvo un valor de  $\kappa$  de 0.597746 ( $\approx 0.600$ ), con una precisión de 0.696277 ( $\approx 0.700$ ); mientras que la lista de 1468 frases con las respuestas válidas de *ChatGPT* (96 % del corpus) obtuvo un valor de  $\kappa$  de 0.420, con una precisión de 0.539, como se presenta en la Tabla 30:

**Tabla 30**

*Comparación del valor de  $\kappa$  para los anotadores humanos y para ChatGPT*

Anotador	Total de frases	Kappa de Fleiss	Precisión ( <i>accuracy</i> )
<b>Humano</b>	1531	0.597746	0.696277
<b>ChatGPT</b>	1468	0.420144	0.538863

El desempeño de *ChatGPT* en la presente tesis muestra valores inferiores a los reportados en los antecedentes, cuando se comparó el desempeño de *ChatGPT* con el de anotadores humanos entrenados o los subcontratados de servicios como MTurk. Huang et al. (2023) afirman que *ChatGPT* identificó correctamente 80 % de los tuits de odio implícito y que el modelo ofreció mejores explicaciones de su clasificación que las de los clasificadores humanos. Por su lado, Gilardi et al. (2023) reportaron que en cuatro de las cinco tareas asignadas a *ChatGPT*, el modelo obtuvo mayor precisión que los anotadores de MTurk con respecto al *gold standard*, y que el acuerdo de *ChatGPT* estuvo entre un 84 - 95 %, dependiendo de la temperatura, mientras que el de los anotadores humanos entrenados estuvo entre 50 y 78.3 %, dependiendo de la tarea y de su dificultad. Sin embargo, Gilardi et al. (2023) no indicaron contra cuántos anotadores humanos se comparó el desempeño de *ChatGPT*, de quienes estos autores afirman que sus respuestas tienden a ser de calidad insuficiente.

Törnberg (2023) comparte resultados similares a los de Gilardi et al. (2023), cuando asevera *ChatGPT-4* superó a los anotadores entrenados y a los de MTurk al identificar la afiliación política en tuits. Törnberg (2023) reportó que *ChatGPT-4* obtuvo mayor precisión y fiabilidad que los anotadores humanos, mientras que presentó igual o menor sesgo o inclinación política que los clasificadores humanos. Por su lado, Zhu et al. (2023) encontraron que *ChatGPT* alcanzó, en general, un nivel de precisión de 0.609 y un F1 score de 0.65 anotando la postura, detección, análisis de sentimientos, incitación al odio y detección de *bots* en tuits.

Las tareas estudiadas en los antecedentes para evaluar el comportamiento y desempeño de *ChatGPT* como potencial clasificador y anotador automático son de una complejidad relativa si se comparan con el nivel de abstracción y de análisis que requiere el etiquetado de ocho roles temáticos para frases sin contexto en español. En la Tabla 30 se demuestra que, el acuerdo entre anotadores humanos con formación lingüística y el de *ChatGPT* con respecto al corpus *gold*

*standard* creado por la investigadora pertenecen ambos al nivel moderado, aunque los anotadores humanos tuvieron mayor precisión en la tarea que *ChatGPT*.

## 2.2 Discusión de los resultados

En esta investigación se creó el *Corpus COVID-19*, el cual contiene los análisis lingüísticos para los tres niveles de las estructuras propuestas por LFG: la estructura-a (estructura argumental) para los roles semánticos; la estructura-f para la representación de las funciones gramaticales (estructura interna); y la estructura-c para la representación de los constituyentes sintácticos (estructura externa). A diferencia de otros estudios similares presentados en los antecedentes, en esta tesis se decidió etiquetar los argumentos y los adjuntos del predicado verbal para incluir en el estudio las oraciones subordinadas que son parte importante del lenguaje natural usado por hablantes reales. No se limitó a diferenciar los argumentos de los complementos circunstanciales o adjuntos únicamente, sino que a estos últimos se les asignó también su función modificadora en la frase y el tipo de sintagma que los representan. Cuando se analizaron las tres estructuras para cada frase, se mantuvo el mismo número de ID en los tres niveles, a fin de que luego se pueda crear un mapeo entre los análisis de estas estructuras. Este proceso de mapeo no se propuso como otro objetivo porque se sale del alcance lingüístico que tiene la presente investigación. Sin embargo, se deja el corpus anotado con la información para futuros trabajos que busquen expandir esta área.

Se trabajó con una lista extensa de roles temáticos para la estructura-a, con la idea de etiquetar el comportamiento de todos los componentes del predicado verbal. Con esta lista se diferenció entre agente y causa, meta y beneficiario, instrumento y compañía, fuente y origen, y tema y paciente, los cuales se fusionaron después para crear el modelo de anotación. La tesaria anotó el *Corpus COVID-19* tomando en cuenta los roles temáticos propuestos en la Gramática Léxico-Funcional y la jerarquía de prominencia de los roles temáticos, los *Mapping Principles* (MP) y la *Wellformedness Condition* (WC), las anotaciones del corpus *SenSem* en *GRIAL*, las etiquetas de *FreeLing/TextServer* y los roles temáticos en inglés de *PropBank* para el análisis de los adjuntos. Con todo este respaldo académico se buscó producir un corpus *gold standard* con el que se pueda trabajar desde otras disciplinas.

La investigadora, además, procuró asignar una etiqueta a cada argumento dependiendo del sentido que el verbo principal presentaba en la frase analizada, sin generalizar sentidos ni emplear

proyecciones de posibles roles; no se limitó tampoco al uso de los proto-roles <agente/paciente>, ampliamente utilizados en la automatización de los roles temáticos y en tareas de etiquetado para NLP. Asimismo, la tesiera mantuvo ordenadas las categorías para cada estructura, dejando las etiquetas de funciones gramaticales dentro de esta estructura y no mezclándolas con la de los roles temáticos, como sucedió en otros trabajos discutidos en el apartado de los antecedentes y en los resultados.

Originalmente, el modelo de anotación se pensó para anotadores humanos, porque, cuando se propuso la investigación ante el Posgrado de Maestría Académica en Lingüística a mediados de 2022, *ChatGPT* no estaba disponible al público. Es por esta razón que se requirió la modificación posterior de la encuesta para poder entrevistar efectivamente al modelo. Cabe recordar que nunca fue parte de los objetivos de esta tesis entrenar *chatbots* ni mucho menos entrenar LLMs como *ChatGPT*; por consiguiente, la entrevista a *ChatGPT* no se realizó aplicando los métodos más comunes para entrenamiento, como el *fine-tuning* o el *prompt-tuning* con contexto (*in-context learning*) o por aprendizaje de varios-tiros (*few-shot learning*). Se prefirió el *zero-shot learning* para que el modelo estuviera en condiciones similares a las de los codificadores humanos y evitar así una comparación tendenciosa entre las anotaciones obtenidas por ambas partes.

Desde finales del 2022, el mundo entero ha sido testigo del giro que ha tomado la aplicación de la inteligencia artificial en la vida cotidiana y, en particular, en la académica. La investigación presentada aquí cuestiona la fiabilidad del desempeño sin entrenamiento de *ChatGPT* en la tarea de clasificar argumentos verbales con ocho posibles roles temáticos para frases en español. El comportamiento que presentó este modelo es moderado, de acuerdo con la estadística que obtuvo, cuando se toma en cuenta que los roles temáticos es uno de los temas más complejos de analizar, principalmente cuando no se tiene entrenamiento avanzado. Los anotadores humanos superaron tanto en el acuerdo como en la precisión a *ChatGPT*, aunque el nivel de acuerdo de los humanos y el del modelo son ambos de nivel moderado. La investigadora considera que se obtendrían mejores resultados si los anotadores recibieran mayor entrenamiento en los análisis sintáctico, gramatical y semántico, principalmente en esta última área lingüística.

*ChatGPT* presentó el comportamiento esperado por la investigadora cuando varió sus respuestas, dependiendo de cómo se le preguntara o de si se le repetía el comando. Sus respuestas también iban decayendo en calidad conforme pasaba el tiempo, porque parecía olvidar qué se le había preguntado cada diez frases, aproximadamente. La disminución en la calidad de sus

respuestas tiene que ver con lo que comúnmente se le ha estado llamando ‘alucinaciones’ del modelo y que, para solucionarlas, se debió repetir el *prompt* múltiples veces, lo cual resultó en un retraso en la entrevista y en este estudio. La investigadora considera que dedicar tiempo a la reiteración de preguntas previene el aprovechamiento de *ChatGPT* como herramienta para etiquetado de los roles temáticos. Como método de consulta, este modelo resulta competitivo e, incluso, atractivo; sin embargo, sus respuestas se deben interpretar con pericia científica: el usuario lego podría tomar por válidas respuestas que no lo son, incluidos quienes pertenecen al espacio académico.

Los valores cuantitativos respaldados con métricas estadísticas pertinentes para evaluar las anotaciones que se reportan aquí, tanto para los codificadores humanos como para *ChatGPT*, reconocen la posibilidad del acuerdo que ocurre por azar entre dos o más anotadores, un asunto bastante probable de que acontezca en este tipo de anotaciones cuando no hay un entrenamiento previo en ninguno de los participantes. Con esta investigación se ofrecen valores estadísticos del coeficiente de kappa de Fleiss que calcularon el acuerdo entre dos o más anotadores por número de anotadores y por número de frase, de manera que las respuestas dadas en otro conjunto de frases por otro conjunto de anotadores no afectaron el valor obtenido para cada caso reportado en los resultados, a no ser que se ofreciera un promedio grupal para fines de representación cuantitativa. El coeficiente empleado aquí tomó en cuenta también la cantidad de categorías que se trabajaron, considerando así todas las posibles variables que podrían afectar la evaluación efectuada.

Finalmente, en esta investigación se consiguió dejar a la academia un corpus anotado por la investigadora, con el cual se pueda entrenar eventualmente sistemas de pregunta-respuesta y modelos como *ChatGPT* con métodos *fine-tuning*. El modelo de anotación creado y evaluado con esta tesis es replicable y, por supuesto, mejorable. Habría sido ideal corregir las anotaciones de desempate junto con los 16 anotadores humanos, para haber enriquecido el corpus con la justificación de los roles temáticos elegidos en la entrevista. No obstante, no hubo espacio para este tipo de reunión y, cuando la autora de este trabajo expuso los resultados del acuerdo ante el grupo de codificadores, estos no mostraron interés en debatirlos.

### CAPÍTULO III. CONCLUSIONES

El motivo de esta investigación era desarrollar una propuesta de modelo de anotación de roles temáticos en español, específico para sistemas de pregunta-respuesta. Para alcanzar este objetivo se creó un corpus anotado con los análisis de las tres estructuras que propone la Gramática Léxico-Funcional: la estructura-c para la representación de los constituyentes sintácticos o estructura externa; la estructura-f para la representación de las funciones gramaticales o estructura interna y la estructura-a o estructura argumental.

Se analizaron estos tres niveles porque la tesiaría consideró oportuno comprender las relaciones sintáctico-semánticas del corpus para poder asignar el rol temático pertinente para cada argumento y adjunto del predicado verbal analizado. Los resultados demuestran que, de no haberse efectuado un estudio exhaustivo de los tres niveles, habría sido difícil etiquetar apropiadamente el corpus y producir, a partir de él, un modelo de anotación sin etiquetas confusas que pertenecieran a otro nivel distinto del semántico, como se ha dado en otras investigaciones.

La Gramática Léxico-Funcional permitió a la tesiaría llevar a cabo análisis ordenados y minuciosos en cada estructura, y establecer categorías claras para cada una de ellas, con jerarquías de prominencia que guiaron a la tesiaría a asignar etiquetas acertadas, por medio de mapeos y justificaciones lógicas. El uso de las plataformas *Sketch Engine*, *GRIAL*, *TextServer* y *FreeLing* brindó una pauta que invitó a la investigadora a cuestionar investigaciones anteriores, y esto le permitió perfeccionar el rumbo que tomaría el etiquetado final del *Corpus COVID-19*. Aun cuando se puede mejorar el modelo presentado con esta tesis, las bases teóricas y el grado crítico con que se abordó esta investigación garantizan haber dejado un trabajo bien elaborado para futuras investigaciones lingüísticas e informáticas.

Gracias al grupo de codificadores fue posible etiquetar completamente el *Corpus COVID-19* y no solamente una parte de este, como es común en estudios con recursos humanos calificados y financieros limitados. La investigadora consideró valioso el aporte humano, porque con sus anotaciones se enriqueció un área de la Lingüística que ha sido poco explotada para el español y que se está convirtiendo cada vez más necesaria con los LLMs. Asimismo, haber contado con la participación de otros estudiantes de lingüística afirmó que se debe dar mayor formación académica en semántica y en análisis gramaticales.

Aparte de esto, el acuerdo entre los anotadores fue más relevante cuantos más codificadores etiquetaron el corpus, particularmente en una tarea tan complicada como clasificar los roles

temáticos. En este aspecto, habría resultado más provechoso haber contado con anotadores que tuvieran mayor entrenamiento en el tema, para que las etiquetas del corpus *gold standard* dejado por la tesiaría se hubieran discutido y corroborado con su aporte académico. No obstante, el propósito de la investigación era desarrollar una propuesta de anotación de los roles temáticos y este objetivo sí se logró exitosamente.

El otro objetivo que redireccionó la investigación fue contar con la participación de *ChatGPT* como herramienta anotadora. Como se mencionó en el apartado de discusiones, el comparar su desempeño en la anotación de roles temáticos en español no era parte de los objetivos originales y, mucho menos, lo era su entrenamiento. Sin embargo, los cambios en el procesamiento de lenguaje natural es una constante que los investigadores deben estar atentos a seguir, porque adaptar los objetivos de cualquier investigación en curso puede modificar enormemente la aplicación de los resultados esperados.

En esta investigación se estudió por primera vez el comportamiento de *ChatGPT* en tareas de NLP, específicamente desde una perspectiva lingüística a nivel académico, con lo que se ha demostrado que hay un considerable espacio en donde el aporte de lingüistas, particularmente aquellos con formación en sintaxis, semántica y en Lingüística Computacional, hará la diferencia en el futuro de NLP. Los corpus lingüísticos anotados automáticamente por *parsers* y sin revisión científica no ofrecen un entrenamiento fiable para los LLMs, como se ha establecido en esta investigación. Es por esta razón que, contar con un corpus anotado, como el desarrollado con esta tesis, permitirá construir herramientas de aprendizaje automático para anotación de roles temáticos, así como que también permitirá evaluar herramientas existentes y comparar la calidad de su anotación.

*ChatGPT* obtuvo un valor de  $\kappa$  de 0.420, con una precisión de 0.539 al anotar el 96 % del *Corpus COVID-19* (1468 frases respondidas), cuando se le compara con el *gold standard* desarrollado por la investigadora. Por su lado, los anotadores humanos etiquetaron el 100 % de las 1531 frases y obtuvieron un valor de  $\kappa$  de 0.597746 ( $\approx 0.600$ ), con una precisión de 0.696277 ( $\approx 0.700$ ). En general, *ChatGPT* obtuvo un rendimiento inferior al de los anotadores humanos, por lo que la investigadora cree que este modelo no está preparado para esta tarea sin haber pasado por una formación lingüística profunda. Con entrenamiento, *ChatGPT* se podría convertir en una herramienta con potencial para clasificar grandes cantidades de datos lingüísticos, pero, por el momento, se requiere más entrenamiento por humanos lingüistas.

Algunas de las herramientas empleadas para desarrollar la clasificación, la anotación y la estadística no fueron las ideales para estas tareas. Excel, por ejemplo, falló en varias ocasiones cuando se buscaba analizar los datos lingüísticos del corpus, reiniciando los documentos a versiones anteriores y retrasando el avance con esto. También falló en mantener el formato dado a las frases y esto causó confusión en un grupo de anotadores y una mayor carga de trabajo para la tesiaría con las correcciones manuales de estos errores. Se sugiere buscar otras opciones que sean más amigables con el procesamiento de grandes cantidades de datos lingüísticos y, quizá, el aporte de otras academias en este punto.

El uso de los *scripts* en *SQL Server*, creados para administrar los datos de las encuestas y para reportar las relaciones entre los anotadores por parte del ingeniero en sistemas informáticos, permitió obtener datos precisos con los cuales se pudiera aplicar la estadística necesaria para evaluar el modelo de anotación y el desempeño de los codificadores. La participación de un experto en el uso de este programa facilitó el proceso enormemente y su invaluable contribución produjo datos respaldados por un cómputo adecuado a estos.

En resumen, se deja a la academia análisis lingüísticos en tres niveles ordenados, con categorías claras y relaciones lógicas entre funciones gramaticales y semánticas, con un estudio exhaustivo de roles temáticos en español y con etiquetas precisas. Además, se produjo un estudio con anotaciones humanas y de modelo de gran tamaño en un área de lingüística poco explotada en el español, necesaria para el futuro de AI. Por ende, esta investigación deja un material que servirá para entrenar a futuros estudiantes de diversas áreas en una rama de la Lingüística todavía abstracta y oscura para muchos, y también para reentrenar *ChatGPT* en la anotación de roles temáticos.

## Bibliografía

- Alcántara-Plá, M. (n.d.). *Lista de corpus*. Inicios. Recuperado el 10 de octubre de 2020, de <https://inicios.es/corpus/>.
- Aljanabi, M., Ghazi, M., Ali, A. H. y Abed, S. A. (2023). ChatGpt: Open Possibilities. *Iraqi Journal For Computer Science and Mathematics*, 4(1), 62-64. Recuperado el 6 de abril de 2023, de [https://scholar.googleusercontent.com/scholar?q=cache:dRA6GTw6T84J:scholar.google.com/+ChatGpt:+Open+Possibilities&hl=es&as\\_sdt=0,5](https://scholar.googleusercontent.com/scholar?q=cache:dRA6GTw6T84J:scholar.google.com/+ChatGpt:+Open+Possibilities&hl=es&as_sdt=0,5).
- Alonso-Cortés, Á. (2004). Donde la teoría computacional de la mente falla. *Revista de libros de la Fundación Caja Madrid*(88), 31-32. Recuperado el 15 de junio de 2022, de <https://dialnet.unirioja.es/servlet/articulo?codigo=836000>.
- Amer, E., Hazem, A., Farouk, O., Louca, A., Mohamed, Y. y Ashraf, M. (2021). A Proposed Chatbot Framework for COVID-19. *2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, pp. 263-268, doi: 10.1109/MIUCC52538.2021.9447652. Recuperado el 15 de junio de 2022, de <https://ieeexplore.ieee.org/abstract/document/9447652>.
- Atserias, J., Climent, S., Farreres, J., Rigau, G. y Rodríguez, H. (2000). Combining multiple methods for the automatic construction of multilingual WordNets. *Amsterdam studies in the theory and history of linguistic science series 4*, 327-340. Recuperado el 15 de junio de 2022, de <https://www.cs.upc.edu/~nlp/papers/atserias97.pdf>.
- Ballesteros, M., Bohnet, B., Mille, S. y Wanner, L. (2016). Data-driven deep-syntactic dependency parsing. *Natural Language Engineering*, 22(6), 939-974. <http://doi.org/10.1017/S1351324915000285>.
- Bel, N. (2010). Platform for Automatic, Normalized Annotation and Cost- Effective Acquisition of Language Resources for Human Language Technologies. PANACEA. *Procesamiento Del Lenguaje Natural*, 45, 327-328. Recuperado de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/826/680>.
- Benites, L. (2021a, 17 de noviembre). *Muestra Bootstrap: Definición, ejemplo*. Statologos. Recuperado el 4 de junio de 2023, de <https://statologos.com/muestra-de-arranque/>.

- Benites, L. (2021b, 23 de noviembre). *Kappa de Fleiss*. Statologos. Recuperado el 5 de junio de 2023, de <https://statologos.com/fleiss-kappa/>.
- Benites, L. (2022, 7 de junio). *Estimación de confiabilidad Alfa de Krippendorff: Definición simple*. Statologos. Recuperado el 04 de junio de 2023, de <https://statologos.com/alfa-de-krippendorff/>.
- Berwick, R. C. (2009). What Genes Can't Learn about Language. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), 1685-1686. <https://doi.org/10.1073/pnas.081287110>.
- Bogantes, P. (2019). "Sayú", el nuevo robot inteligente de Correos de Costa Rica. Sinart. costaricamedios.cr. Recuperado el 30 de abril de 2022, de <https://costaricamedios.cr/2019/12/20/sayu-el-nuevo-robot-inteligente-de-correos-de-costa-rica/>.
- Bonales Daimiel, G. y Martínez Estrella, E. C. (2021). Using Virtual Assistants and Chatbots for Crisis Communication. *aDResearch: Revista Internacional de Investigación en Comunicación*, (25), 70-91. Recuperado el 15 de junio de 2022, de <https://dialnet.unirioja.es/servlet/articulo?codigo=7870939>.
- Bonial, C., Hwang, J., Bonn, J., Conger, K., Babko-Malaya, O., y Palmer, M. (2015). English propbank annotation guidelines. *Center for Computational Language and Education Research Institute of Cognitive Science University of Colorado at Boulder*, 48. Recuperado el 8 de junio de 2023, de <https://verbs.colorado.edu/propbank/EPB-Annotation-Guidelines.pdf>.
- Bosque, I. y Gutiérrez-Rexach, J. (2009). *Fundamentos de sintaxis formal*. Ediciones Akal, S.A., Madrid, España.
- Bravo-Segal, S. y Villar, F. (2020). La representación de los mayores en los medios durante la pandemia COVID-19: ¿hacia un refuerzo del edadismo?. *Rev Esp Geriatr Gerontol*. <https://doi.org/10.1016/j.regg.2020.06.002>.
- Bresnan, J. (2001). *Lexical-functional syntax*. Blackwell Publishers Ltd., Oxford, UK.
- Bresnan, J., Asudeh, A., Toivonen, I. y Wechsler, S. (2016). *Lexical-Functional Syntax*. (2da ed.). Blackwell Publishers Ltd., West Sussex, UK.

- Brosa Rodríguez, A. (2022). Cambios Semánticos en virus Y pandemia desde la Lingüística de Corpus y la sociolingüística. CEEOL. Recuperado el 9 de mayo de 2023, de <https://www.ceeol.com/search/chapter-detail?id=1069928>.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... y Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901. <https://doi.org/10.48550/arXiv.2005.14165>.
- Butt, M. (2006). *Theories of Case*. Cambridge University Press.
- Calvo, H. y Gelbukh, A. (2008). Determinación automática de roles semánticos usando preferencias de selección sobre corpus muy grandes. *Computación y Sistemas*, 12(1), 128-150. Recuperado el 15 de junio de 2022, de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-55462008000300008&lng=es&tlng=](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462008000300008&lng=es&tlng=).
- Canete, J., Chaperon, G., Fuentes, R., Ho, J. H., Kang, H. y Pérez, J. (2020). Spanish pre-trained bert model and evaluation data. *Pml4dc at iclr, 2020*, 1-10. Recuperado el 20 de julio de 2022, de <https://users.dcc.uchile.cl/~jperez/>.
- Cao, Z., Yang, Y., & Zhao, H. (2023). AutoHall: Automated Hallucination Dataset Generation for Large Language Models. *arXiv preprint arXiv:2310.00259*. <https://doi.org/10.48550/arXiv.2310.00259>.
- Cascón-Katchadourian, J. (2020). Tecnologías para luchar contra la pandemia Covid-19: Geolocalización, rastreo, big data, SIG, inteligencia artificial y privacidad // Technologies to fight the Covid-19 pandemic: Geolocation, tracking, big data, GIS, artificial intelligence, and privacy. *Professional de la información*, v.29, n.4, e290429. <https://doi.org/10.3145/epi.2020.jul.29>.
- Castellón, I., Climent, S., Coll-Florit, M., Lloberes, M. y Rigau, G. (2012). Constitución de un corpus de semántica verbal del español: Metodología de anotación de núcleos argumentales. *RLA. Revista de lingüística teórica y aplicada*, 50 (1), 13-38. <https://dx.doi.org/10.4067/S0718-48832012000100002>.
- Chan, C., Cheng, J., Wang, W., Jiang, Y., Fang, T., Liu, X., & Song, Y. (2023). Chatgpt evaluation on sentence level relations: A focus on temporal, causal, and discourse relations. *arXiv preprint arXiv:2304.14827*. <https://doi.org/10.48550/arXiv.2304.14827>.

- ChatGPT. (2023). Respuesta generada a la pregunta: ¿Podrías definir para mí el acuerdo entre anotadores en el procesamiento del lenguaje natural? [Mensaje en un foro de discusión en línea]. Recuperado de <https://chat.openai.com/share/eff6e442-c12c-4f07-bdcb-b73f5018ddfc>.
- Chen, Q., Leaman, R., Allot, A., Luo, L., Wei, C.-H., Yan, S. y Lu, Z. (2021). Artificial Intelligence in Action: Addressing the COVID-19 Pandemic with Natural Language Processing. *Annual Review of Biomedical Data Science*, 4(1), 313-339. <https://doi.org/10.1146/annurev-biodatasci-021821-061045>.
- Chiarcos, C., Ionov, M., Glaser, L. y Fäth, C. (2021). An Ontology for CoNLL-RDF: Formal Data Structures for TSV Formats in Language Technology. In *3rd Conference on Language, Data and Knowledge (LDK 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. Recuperado el 15 de junio de 2022, de <https://drops.dagstuhl.de/opus/volltexte/2021/14556/>.
- Chomsky, N. (1987). *Estructuras sintácticas* (C. Peregrín Otero, Trad.). México, España, Bogotá, Argentina: Siglo Veintiuno.
- Chomsky, N., Roberts, I. y Watumull, J. (8 de marzo de 2023). *Noam Chomsky: The false promise of chatgpt*. The New York Times. Recuperado el 8 de abril de 2023, de <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>.
- Colic, N., Furrer, L. y Rinaldi, F. (29 de junio de 2020). *Annotating the Pandemic: Named Entity Recognition and Normalisation in COVID-19 Literature*. ACL 2020 Workshop on Natural Language Processing for COVID-19 (NLP-COVID). Recuperado el 4 de octubre, 2020, de <https://openreview.net/group?id=aclweb.org/ACL/2020/Workshop/NLP-COVID>.
- Colic, N. y Rinaldi, F. (2019). Improving spaCy dependency annotation and PoS tagging web service using independent NER services. *Genomics & Informatics*, 17(2). Recuperado el 15 de junio de 2022, de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6808626/>.
- Comisión Económica para América y el Caribe (CEPAL), Naciones Unidas. (7 de abril de 2020). Las oportunidades de la digitalización en América Latina frente al COVID-19. Recuperado el 16 de setiembre de 2021, de <https://www.cepal.org/es/publicaciones/45360-opportunidades-la-digitalizacion-america-latina-frente-al-covid-19>.
- Coniam, D. (2014). The linguistic accuracy of chatbots: usability from an ESL perspective. *Text & Talk*, 34(5), 545-567. <https://doi.org/10.1515/text-2014-0018>.

- Cothenet, C. (2019). *Dependency Parser or how to find syntactic neighbours of a word*. Medium. Recuperado el 16 de julio de 2022, de <https://towardsdatascience.com/dependency-parser-or-how-to-find-syntactic-neighbours-of-a-word-a9e7d17ffe8>.
- Culicover, P. W. y Jackendoff, R. (2005). *Simpler syntax*. Oxford University Press. United Kingdom.
- De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E. y Rizzo, C. (2023). ChatGPT and the Rise of Large Language Models: The New AI-Driven Infodemic Threat in Public Health. *SSRN 4352931*. <https://dx.doi.org/10.2139/ssrn.4352931>.
- Degrauwe, J. y Goethals, P. (2020). Reflexive pronouns in Spanish universal dependencies. *Procesamiento del Lenguaje Natural*, 64, 77-84. Recuperado el 15 de junio de 2022, de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6198>.
- Dehaene, S. (2015). *El cerebro lector: Últimas noticias de las neurociencias sobre la lectura, la enseñanza, el aprendizaje y la dislexia* (M. J. D'alessio, Trans.; 1a ed.). Siglo Veintiuno Editores.
- Deng, J. y Lin, Y. (2022). The Benefits and Challenges of ChatGPT: An Overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. <https://doi.org/10.54097/fcis.v2i2.4465>.
- Deng, S. (2023). *Conll*. CoNLL. Recuperado el 19 de mayo de 2023, de <https://www.conll.org/>.
- Departamento de Ciencias de la Computación Universidad de Chile. (2022). *GitHub - dccuchile/beto: BETO - Spanish version of the BERT model*. GitHub. Recuperado el 14 de julio de 2022, de <https://github.com/dccuchile/beto>.
- Devlin, J., Chang, M., Lee, K. y Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv.org. Recuperado el 14 de julio de 2022, de <https://arxiv.org/abs/1810.04805>.
- Devlin, J. y Chang, M. (2018). *Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing*. Google AI Blog. Recuperado el 14 de julio de 2022, de <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>.
- Drăgușin, D. (2015). Thematic Relations in Syntax. *Intertext*, 35(3-4), 55-68. ISSN 1857-3711. Recuperado el 13 de junio de 2023, de [https://ibn.idsi.md/vizualizare\\_articol/44727](https://ibn.idsi.md/vizualizare_articol/44727).
- Dueñas Luna, G. E. y González, F. A. (30 de julio de 2012). Anotación y descripción de textos digitales sin formato de la base de casos médicos de la Facultad de Medicina de la Universidad Nacional de Colombia. *Cuadernos de Lingüística Hispánica*, (20), 83-98.

- Recuperado el 15 de junio de 2022, de [https://revistas.uptc.edu.co/index.php/linguistica\\_hispanica/article/view/462](https://revistas.uptc.edu.co/index.php/linguistica_hispanica/article/view/462).
- Eloundou, T., Manning, S., Mishkin, P. y Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*. Recuperado el 6 de abril de 2023, de <https://arxiv.org/abs/2303.10130#>.
- Esteva, A., Kale, A., Paulus, R., Hashimoto, K., Yin, W., Radev, D. y Socher, R. (2020). Co-search: Covid-19 information retrieval with semantic search, question answering, and abstractive summarization. *arXiv preprint arXiv:2006.09595*. Recuperado el 4 de octubre, 2020, de <https://arxiv.org/pdf/2006.09595.pdf>.
- Farlex. (s.f.). *Diccionario de Español / spanish dictionary*. The Free Dictionary. Recuperado el 13 de abril de 2023, de <https://es.thefreedictionary.com/>.
- Fernández Moreno, D. (2021). Los *chatbots* se convierten en el aliado perfecto para el contact center. *Tecnohotel: revista profesional para la hostelería y restauración*, (487), 34-35. Recuperado el 15 de junio de 2022, de <https://dialnet.unirioja.es/servlet/articulo?codigo=7781916>.
- Fernández, A y Vázquez, G. (2015). SenSem Databank LDC2015T02. Web Download. Philadelphia: Linguistic Data Consortium. Recuperado el 15 de junio de 2022, de <https://catalog.ldc.upenn.edu/LDC2015T02>.
- Fichtel, L., Kalo, J. C., y Balke, W. T. (2021). Prompt tuning or fine-tuning-investigating relational knowledge in pre-trained language models. In *3rd Conference on Automated Knowledge Base Construction*. Recuperado el 24 de mayo de 2023, de <https://openreview.net/forum?id=o7sMlpr9yBW>.
- Fodor, J. A. (1986). *La modularidad de la mente: un ensayo sobre la psicología de las facultades* (J. M. Igoa, Trans.). Morata. (The modularity of the mind).
- Gallagher, S., Rafferty, A. y Wu, A. (2004). *Natural Language Processing*. NLP-Overview. Recuperado el 9 de enero de 2021, de [https://cs.stanford.edu/people/eroberts/courses/soco/projects/2004-05/nlp/overview\\_history.html](https://cs.stanford.edu/people/eroberts/courses/soco/projects/2004-05/nlp/overview_history.html).
- Garbade, M. J. (15 de octubre de 2018). A Simple Introduction to Natural Language Processing. *Medium, Becoming Human: Artificial Intelligence Magazine*, Recuperado el 15 de junio

- de 2022, de <https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32>.
- García Marchena, O. (2009). Lingüística española e Inteligencia Artificial Aplicación informática de gramáticas de restricciones para la confección de agentes de diálogo. *Interlingüística*, (18), 472-483. Recuperado el 15 de junio de 2022, de <https://dialnet.unirioja.es/servlet/articulo?codigo=3130252>.
- Gazdar, G., Klein, E., Pullum, G. K. y Sag, I. A. (1985). *Generalized phrase structure grammar*. Harvard University Press.
- Gelbukh, A. (2010). Procesamiento del lenguaje natural y sus aplicaciones. *Komputer Sapiens*, 1, 6-11. Recuperado el 15 de junio de 2022, de [https://scholar.googleusercontent.com/scholar?q=cache:OcdugTvr7Y0J:scholar.google.com/+Procesamiento+del+lenguaje+natural+y+sus+aplicaciones&hl=es&as\\_sdt=0,5](https://scholar.googleusercontent.com/scholar?q=cache:OcdugTvr7Y0J:scholar.google.com/+Procesamiento+del+lenguaje+natural+y+sus+aplicaciones&hl=es&as_sdt=0,5).
- George, A. S., George, A. H. y Martin, A. G. (2023). A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*, 1(1), 9-23. <https://doi.org/10.5281/zenodo.7644359>.
- Gerdes, K. y Kahane, S. (2016). Dependency Annotation Choices: Assessing Theoretical and Practical Issues of Universal Dependencies. In *Proceedings of the 10th Linguistic Annotation Workshop held in conjunction with ACL 2016 (LAW-X 2016)* (pp. 131-140). <https://halshs.archives-ouvertes.fr/halshs-01509118/>.
- Gilardi, F., Alizadeh, M. y Kubli, M. (2023). Chatgpt outperforms crowd-workers for text-annotation tasks. *arXiv preprint arXiv:2303.15056*. <https://doi.org/10.48550/arXiv.2303.15056>.
- Gil-Vallejo, L., Castellón, I., Coll-Florit, M. y Turmo, J. (2015). Hacia una clasificación verbal automática para el español: estudio sobre la relevancia de los diferentes tipos y configuraciones de información sintáctico-semántica. *Linguamática*, 7(1), 41-52. Recuperado el 15 de junio de 2022, de <https://upcommons.upc.edu/handle/2117/77807>.
- Grup de Recerca Interuniversitari en Aplicacions Lingüístiques. (s.f.). *GRIAL Léxico verbal del corpus SenSem*. Grial.edu.es. Recuperado el 29 de mayo de 2022, de <http://grial.edu.es/sensem/lexico?idioma=es>.
- Grupo de investigación ELCI. (s.f.). Corpus Iberia. Recuperado el 1 de noviembre de 2020, de: <http://www.investigacion.cchs.csic.es/elci/node/8>.

- Guo, X., Mirzaalian, H., Sabir, E., Jaiswal, A. y Abd-Almageed, W. (2020). COVID-19 Semantic Textual Similarity Dataset. *arXiv preprint arXiv:2007.02461*. <https://doi.org/10.48550/arXiv.2007.02461>.
- Gupta, A., Hathwar, D. y Vijayakumar, A. (2020). Introduction to AI chatbots. *International Journal of Engineering Research and Technology*, 9(7), 255-258. Recuperado el 15 de junio de 2022, de <https://www.ijert.org/introduction-to-ai-chatbots>.
- Gupta, M. (28 de enero de 2019). NLP: Synsets for a word in WordNet. *GeeksforGeeks*. Recuperado el 13 de noviembre de 2021, de <https://www.geeksforgeeks.org/nlp-synsets-for-a-word-in-wordnet/>.
- Haleem, A., Javaid, M. y Singh, R. P. (2023). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil transactions on benchmarks, standards and evaluations*, 100089. <https://doi.org/10.1016/j.tbench.2023.100089>.
- Hariri, W. (2023). Unlocking the Potential of ChatGPT: A Comprehensive Exploration of its Applications, Advantages, Limitations, and Future Directions in Natural Language Processing. *arXiv preprint arXiv:2304.02017*. <https://doi.org/10.48550/arXiv.2304.02017>.
- Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., ... y Awadalla, H. H. (2023). How good are gpt models at machine translation? a comprehensive evaluation. *arXiv preprint arXiv:2302.09210*. <https://doi.org/10.48550/arXiv.2302.09210>.
- Herriman, M., Meer, E., Roy Rosin, M., Vivian Lee, M., Vindell Washington, M. y Kevin G. Volpp, M. (18 de junio de 2020). Asked and Answered: Building a Chatbot to Address COVID-19-Related Concerns. *NEJM Catalyst. Innovations in Care Delivery*. DOI: 10.1056/CAT.20.0230. Recuperado de <https://catalyst.nejm.org/doi/full/10.1056/cat.20.0230>.
- Hinze, A., Heese, R., Schlegel, A. y Paschke, A. (2019). Manual semantic annotations: User evaluation of interface and interaction designs. *Journal of Web Semantics*, 58, 100516. <https://doi.org/10.1016/j.websem.2019.100516>.
- Hodges, A. (n.d.). *The Alan Turing Internet Scrapbook*. Alan Turing Internet Scrapbook The Turing Test. Recuperado el 10 de enero de 2021, de <https://www.turing.org.uk/scrapbook/test.html>.

- Horev, R. (2018). *BERT Explained: State of the art language model for NLP*. Medium. Recuperado el 14 de julio de 2022, de <https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>.
- Huang, F., Kwak, H. y An, J. (2023). Is chatgpt better than human annotators? potential and limitations of chatgpt in explaining implicit hate speech. *arXiv preprint arXiv:2302.07736*. <https://doi.org/10.48550/arXiv.2302.07736>.
- Instituto de Ingeniería del Conocimiento. (n.d.). Procesamiento del Lenguaje Natural Las máquinas analizan nuestro lenguaje en texto. Procesamiento del Lenguaje Natural. Recuperado el 7 de enero de 2021, de <https://www.iic.uam.es/inteligencia-artificial/procesamiento-del-lenguaje-natural/>.
- Jara Murillo, C. V (2013). El treebank del español" IPROCOLDI": componente anotado del corpus CODIMEP-CR. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 39(2), 143-171. Recuperado el 15 de junio de 2022, de <https://revistas.ucr.ac.cr/index.php/filyling/article/view/15094>.
- Jasarevic, T., Adhanom Ghebreyesus, T., Ryan, M., Van Kerkhove, M. y Solomon, S. (17 de abril de 2020). *COVID-19 virtual press conference*. [Discurso principal]. Conferencia de la Organización Mundial de la Salud, Ginebra, Suiza. Recuperado el 28 de agosto de 2021, de [https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencias-coronavirus-press-conference-17apr2020.pdf?sfvrsn=1bd9c4cd\\_2](https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencias-coronavirus-press-conference-17apr2020.pdf?sfvrsn=1bd9c4cd_2).
- Joan Casademont, A. (2020). Los verbos fraseológicos especializados en el ámbito de los museos. *Revista De Lenguas Modernas*, (33), 85–109. <https://doi.org/10.15517/rlm.v0i33.38912>.
- Johnson. (26 de abril de 2023). *CHATGPT raises questions about how humans acquire language*. The Economist. Recuperado el 21 de mayo de 2023, de <https://www.economist.com/culture/2023/04/26/chatgpt-raises-questions-about-how-humans-acquire-language>.
- Kaplan, R. M. y Bresnan, J. (1982). Lexical-Functional Grammar: a formal system for grammatical representation. En Joan Bresnan (ed.): *The mental representation of grammatical relations*, Cambridge (Mass.): MIT Press, 173-280.

- Kleinberg, B., van der Vegt, I. y Mozes, M. (2020). Measuring emotions in the covid-19 real world worry dataset. *arXivpreprint arXiv:2004.04225*.  
<https://doi.org/10.48550/arXiv.2004.0422>.
- Kocoń, J., Cichecki, I., Kaszyca, O., Kochanek, M., Szydło, D., Baran, J., ... y Kazienko, P. (2023). Chatgpt: Jack of all trades, master of none. *arXiv preprint arXiv:2302.10724*.
- Korstanje, J. (2021). *The F1 score*. Medium. Recuperado el 16 de julio de 2022, de <https://towardsdatascience.com/the-f1-score-bec2bbc38aa6>.
- Koubaa, A., Boulila, W., Ghouti, L., Alzahem, A. y Latif, S. (2023). Exploring ChatGPT Capabilities and Limitations: A Critical Review of the NLP Game Changer. *Preprints.org*, 2023030438. <https://doi.org/10.20944/preprints202303.0438.v1>.
- Koza Orellana, W. (2015). Propuesta de extracción automática de candidatos a término del dominio médico procesando información lingüística. Descripción y evaluación de resultados. *Alfa: Revista de Linguística (São Josédo Rio Preto)*, 59 (1), 113-128. <https://doi.org/10.1590/1981-5794-1502-5>.
- Kuzman, T., Ljubešić, N. y Mozetič, I. (2023). ChatGPT: Beginning of an End of Manual Annotation? Use Case of Automatic Genre Identification. *arXiv preprint arXiv:2303.03953*. <https://doi.org/10.48550/arXiv.2303.03953>.
- La Nación. (2 de julio de 2021). *Covid-19: Cuántos Días Tardan en aparecer los síntomas de Coronavirus*. LA NACION. Recuperado el 24 de abril de 2022, de <https://www.lanacion.com.ar/lifestyle/coronavirus-cuantos-dias-tardan-en-aparecer-los-sintomas-de-covid-19-nid01042021/>.
- Laerd Statistics (2019). Fleiss' kappa using SPSS Statistics. *Statistical tutorials and software guides*. Recuperado el 23 de abril de 2023, de <https://statistics.laerd.com/>.
- Lee, A. (20 de abril de 2023). *What are large language models and why are they important?*. NVIDIA Blog. Recuperado el 21 de mayo de 2023, de <https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-models-used-for/>.
- Lee, E. (2023). Is ChatGPT a False Promise?. United States of America. Recuperado el 8 de abril de 2023, de <https://policycommons.net/artifacts/3528039/is-chatgpt-a-false-promise/4328874/>. CID: 20.500.12592/hzdz6k.
- Lexical Computing Ltd. (s. f.). Sketch Engine [en línea]. Disponible en: <https://www.sketchengine.eu/>.

- Lexical Computing. (2020). esTenTen – Spanish corpus from the web. Sketch Engine. Recuperado el 10 de octubre de 2020, de <https://www.sketchengine.eu/estenten-spanish-corpus/>.
- Li, Z. (2023). The dark side of chatgpt: Legal and ethical challenges from stochastic parrots and hallucination. *arXiv preprint arXiv:2304.14347*. <https://doi.org/10.48550/arXiv.2304.14347>.
- Llisterri, J. (2003). Lingüística y tecnologías del lenguaje. Lynx. *Panorámica de Estudios Lingüísticos (Departament de Teoria dels Llenguatges, Universitat de València) 2*: 9-71. Recuperado el 15 de junio de 2022, de [http://www.iling.unam.mx/cursocorpus/TecnoLing\\_Lynx02.pdf](http://www.iling.unam.mx/cursocorpus/TecnoLing_Lynx02.pdf).
- López Rodríguez, C. I. (2020). Marcos predicativos asociados al concepto signo y síntoma en textos sobre medicina en español. *Revista signos*, 53 (103), 392-418. <https://dx.doi.org/10.4067/S0718-09342020000200392>.
- Lund, B. D. y Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library Hi Tech News*, 40(3), 26-29. <https://doi.org/10.1108/LHTN-01-2023-0009>
- Luque, F. M. (2017). El Lenguaje Natural como Lenguaje Formal. *arXiv preprint arXiv:1703.04417*. <https://doi.org/10.48550/arXiv.1703.04417>.
- Lybarger, K., Ostendorf, M., Thompson, M. y Yetisgen, M. (9 de julio de 2020). Extracting COVID-19 diagnoses and symptoms from clinical text: A new annotated corpus and neural event extraction framework. *Journal of Biomedical Informatics*, 117, 103761. <https://doi.org/10.1016/j.jbi.2021.103761>.
- Manning, C. D. (2022). Human language understanding & reasoning. *Daedalus*, 151(2), 127-138.
- Manning, C. D., Raghavan, P. y Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., Taylor, A. (1999). Treebank-3 LDC99T42. Web Download. Philadelphia: Linguistic Data Consortium. Recuperado el 15 de junio de 2022, de <https://catalog.ldc.upenn.edu/LDC99T42>.
- Marimon, M. (2015). Tibidabo: a syntactically and semantically annotated corpus of Spanish. *Corpora*, 10(3), 259-276. Recuperado el 15 de junio de 2022, de <https://www.eupublishing.com/doi/abs/10.3366/cor.2015.0077>.

- Marimon, M., Bel, N., Fisas, B., Arias, B., Vázquez, S., Vivaldi, J., Morell, C., Lorente, M. (2014). The IULA Spanish LSP Treebank: Building and browsing. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. Recuperado el 15 de junio de 2022, de [https://www.academia.edu/33699111/The\\_IULA\\_Spanish\\_LSP\\_Treebank\\_building\\_and\\_browsing](https://www.academia.edu/33699111/The_IULA_Spanish_LSP_Treebank_building_and_browsing).
- Marimon, M., Fisas Elizalde, B., Bel Rafecas, N., Arias Badia, B., Vázquez, S., Vivaldi, J., ... y Lorente Casafont, M. (2012). The IULA treebank. In *Calzolari N, Choukri K, Declerck T, Uğur Doğan M, Maegaard B, Mariani J, Moreno A, Odijk J, Piperidis S, editors. Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'12); 2012 May 21-27; Istanbul, Turkey. Paris: European Language Resources Association (ELRA); 2012. p. 1920-26. ELRA (European Language Resources Association)*. Recuperado el 15 de junio de 2022, de <https://repositori.upf.edu/handle/10230/46287>.
- Marimon, M. y Bel, N. (2015). Dependency structure annotation in the IULA Spanish LSP Treebank. *Language Resources y Evaluation*, 49(2), 433-454. doi:10.1007/s10579-014-9280-5. <https://doi.org/10.1007/s10579-014-9280-5>.
- Marín Raventós, G. y Sandí Salazar, M. (2017). *Automatización del análisis sintáctico para el español americano con el fin de crear un Treebank estandarizado*. (Tesis de Maestría Académica en Computación e Informática). Universidad de Costa Rica, San José, Costa Rica. Recuperado el 15 de junio de 2022, de <https://revistas.ucr.ac.cr/index.php/kanina/article/view/30232>.
- Márquez Díaz, J. E. (2020). Inteligencia Artificial y Big Data como soluciones frente al COVID-19. *Revista de Bioética y Derecho*, (50), 315-331. Recuperado el 15 de junio de 2022, de [https://scielo.isciii.es/scielo.php?pid=S1886-58872020000300019&script=sci\\_arttext&tlng=pt](https://scielo.isciii.es/scielo.php?pid=S1886-58872020000300019&script=sci_arttext&tlng=pt).
- Martí Antonín, M. A., Taulé Delor, M., Márquez i Villodre, L. y Bertran, M. (2007). Anotación semiautomática con papeles temáticos de los corpus CESS-ECE. *Procesamiento del Lenguaje Natural*, (38), 67-76. Recuperado el 15 de junio de 2022, de <https://www.redalyc.org/pdf/5157/515751738009.pdf>.

- Martínez, M. (2019). *Google BERT: actualización para entender el lenguaje natural - Human Level. Human Level*. Retrieved 15 July 2022, from <https://www.humanlevel.com/articulos/posicionamiento-natural-buscadores/google-bert-actualizacion-para-entender-el-lenguaje-natural.html>.
- Medina, S. y Turmo, J. (2018). Clasificación conjunta de frases clave y sus relaciones en documentos electrónicos de salud en español (Joint classification of Key-Phrases and Relations in Electronic Health Documents). In *TASS@ SEPLN* (pp. 83-88). Recuperado el 15 de junio de 2022, de <https://core.ac.uk/download/pdf/185527962.pdf>.
- Miner, A. S., Laranjo, L. y Kocaballi, A. B. (2020). Chatbots in the fight against the COVID-19 pandemic. *npj Digital Medicine*, 3(65), 1-4; <https://doi.org/10.1038/s41746-020-0280-0>.
- Moeller, S. y Hulden, M. (marzo de 2021). Integrating Automated Segmentation and Glossing into Documentary and Descriptive Linguistics. In *Proceedings of the Workshop on Computational Methods for Endangered Languages* (Vol. 1, pp. 86-95). Recuperado el 19 de mayo de 2023, de <https://journals.colorado.edu/index.php/computel/article/view/965/895>.
- Moreda, P. (2008). *Los roles semánticos en la tecnología del lenguaje humano: anotación y aplicación*. Universidad de Alicante. Recuperado el 15 de junio de 2022, de <http://rua.ua.es/dspace/handle/10045/9598>.
- Moreno-Sandoval, A. y Campillos-Llanos, L. (2013). Design and Annotation of MultiMedica–A Multilingual Text Corpus of the Biomedical Domain. *Procedia-Social and Behavioral Sciences*, 95, 33-39. <https://doi.org/10.1016/j.sbspro.2013.10.619>.
- Müller, M., Salathé, M. y Kummervold, P. E. (2023). COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter. *Frontiers in artificial intelligence*, 6, 1023281. <https://doi.org/10.3389/frai.2023.1023281>.
- Navarro-Colorado, F. B. (2007). *Metodología, construcción y explotación de corpus anotados semántica y anafóricamente* (Tesis doctoral). Universitat d'Alacant-Universidad de Alicante. España. Recuperado el 15 de junio de 2022, de [https://rua.ua.es/dspace/bitstream/10045/7736/1/tesis\\_doctoral\\_francisco\\_de\\_borja.pdf](https://rua.ua.es/dspace/bitstream/10045/7736/1/tesis_doctoral_francisco_de_borja.pdf).
- Neumeister, S. (2020). *Testing of a usability assessment tool for chatbots: investigating the effect of believing that a chatbot might be a human*. University of Twente. Bachelor thesis, The Netherlands. Recuperado el 15 de junio de 2022, de <http://essay.utwente.nl/81714/>.

- Newman, J. y Cox, C. (2020). Corpus Annotation. In M. Paquot y S. T. Gries (Eds.), *A Practical Handbook of Corpus Linguistics* (pp. 25-48). Springer International Publishing. [https://doi.org/10.1007/978-3-030-46216-1\\_2](https://doi.org/10.1007/978-3-030-46216-1_2).
- Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., . . . Zeman, D. (2020). Universal dependencies v2: An ever growing multilingual treebank collection. *arXiv preprint arXiv:2004.10643*. <https://doi.org/10.48550/arXiv.2004.10643>.
- Obando, J., Pulido, J. y Gómez, J. (2020). Procesamiento del Lenguaje Natural para Reconocer Mensajes de Textos Extorsivos a través del Análisis Sintáctico y Lematización. *Revista CIENCIA Y TECNOLOGÍA*, 16(1), 33-42. Recuperado el 15 de junio de 2022, de <https://revistas.unitru.edu.pe/index.php/PGM/article/view/2752>.
- Oh, B.-D., y Schuler, W. (2023). Why Does Surprisal From Larger Transformer-Based Language Models Provide a Poorer Fit to Human Reading Times? *Transactions of the Association for Computational Linguistics*, 11, 336-350. [https://doi.org/10.1162/tacl\\_a\\_00548](https://doi.org/10.1162/tacl_a_00548)
- Oliver, A. y Climent, S. (2011). Construcción de los WordNets 3.0 para castellano y catalán mediante traducción automática de corpus anotados semánticamente. *Procesamiento del lenguaje natural*, 47, 293-300. Recuperado el 15 de junio de 2022, de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/960>.
- OpenAI (2023). GPT-4 Technical Report. *ArXiv*, *abs/2303.08774*. Recuperado el 6 de abril de 2023, de <https://www.semanticscholar.org/paper/GPT-4-Technical-Report-OpenAI/8ca62fdf4c276ea3052dc96dcfd8ee96ca425a48>.
- Schulman, J., Zoph, B., Kim, C., Hilton, J., Menick, J., Weng, J., Ceron Uribe, J.F., Fedus, L., Metz, L., Pokorny, M., Gontijo Lopes, R., Zhao, S., Vijayvergiya, A., Sigler, E., Perelman, A., Voss, C., Heaton, M., Parish, J., Cummings, D., Nayak, R.,... Balcom, V. (2022). OpenAI: Introducing ChatGPT. Recuperado el 6 de abril de 2023, de <https://openai.com/blog/chatgpt>.
- Organización Mundial de la Salud. (2021). *Información básica sobre la COVID-19*. Who.int. Recuperado el 30 de abril de 2022, de <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>.
- Organización Mundial de la Salud. (2022). *Nuevo coronavirus 2019*. Who.int. Recuperado el 2 de mayo de 2022, de <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019>.

- Oyebode, O., Ndulue, C., Adib, A., Mulchandani, D., Suruliraj, B., Orji, F. A., . . . Orji, R. (23 de julio de 2020). Health, Psychosocial, and Social issues emanating from COVID-19 pandemic based on Social Media Comments using Natural Language Processing. *JMIR medical informatics*, 9(4), e22734. <https://doi.org/10.2196/22734>
- Oyebode, O., Ndulue, C., Mulchandani, D., Suruliraj, B., Adib, A., Orji, F. A., . . . Orji, R. (23 de agosto de 2020). COVID-19 Pandemic: Identifying Key Issues using Social Media and Natural Language Processing. *Journal of Healthcare Informatics Research*, 6(2), 174-207. <https://doi.org/10.1007/s41666-021-00111-w>.
- Ozkural, E. (27 de diciembre de 2017). Natural Language Processing (NLP): Chomsky's Theories of Syntax. *eHealth First*. Recuperado el 12 de enero de 2021, de <https://medium.com/@ehfirst/natural-language-processing-nlp-chomskys-theories-of-syntax-92fb8fa3d035>.
- Padró Cicera, L. (2022). *TextServer - linguistic analysis of texts | TALP :: Language and Speech Technologies and Applications*. Talp.upc.edu. Recuperado el 29 de mayo de 2022, de <https://www.talp.upc.edu/demo-detail/438/TextServer>.
- Parodi, G. (2006). El Grial: Interfaz computacional para anotación e interrogación de corpus en español. *RLA. Revista de lingüística teórica y aplicada*, 44(2), 91-115. Recuperado el 15 de junio de 2022, de <https://scielo.conicyt.cl/pdf/rla/v44n2/art07.pdf>.
- Parodi, G. (2010). *Lingüística de corpus: de la teoría a la empiria*. Madrid, Spain: Editorial Iberoamericana / Vervuert. Recuperado el 4 de octubre de 2020, de <https://elibro-net.ezproxy.sibdi.ucr.ac.cr/es/ereader/sibdi/36857?page=37>.
- Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., ... y Tao, D. (2023). Towards making the most of ChatGPT for machine translation. *arXiv preprint arXiv:2303.13780*. <https://doi.org/10.48550/arXiv.2303.13780>.
- Perdigão, F. (17 de setiembre de 2020). How Important is Linguistics in Your Chatbot? *Visor.Ai*. Recuperado el 18 de agosto de 2021, de <https://www.visor.ai/linguistics-in-your-chatbot/>.
- Pérez García, D. (2021). *Chatbots, el futuro hecho realidad*. Three Points. Recuperado el 13 de julio de 2022, de <https://www.threepoints.com/blog/chatbots-el-futuro-hecho-realidad>.
- Periñán Pascual, J. C. (2012). En defensa del procesamiento del lenguaje natural fundamentado en la lingüística teórica. *Onomázein: Revista de Lingüística, Filología y Traducción*, (26), 13-48. <https://doi.org/10.7764/onomazein.26.01>.

- Peris Morant, A. y Taulé Delor, M. (2011). AnCora-Nom: A Spanish Lexicon of Deverbal Nominalizations. *Procesamiento del Lenguaje Natural*, 46, 11-18. Recuperado el 15 de junio de 2022, de <http://diposit.ub.edu/dspace/handle/2445/128934>.
- Petro, A. (1996). The Language Instinct by Steven Pinker. *TESOL Quarterly*, 30(2), 365-366. <https://doi.org/10.2307/3588154>.
- Pinker, S. (2009). *El instinto del lenguaje: cómo crea el lenguaje la mente* (J. M. Igoa González, Trans.; Sexta ed.). Alianza.
- Pollard, C. y Sag, I. I (1987). *Information-based syntax and semantics*. Stanford: CSLI.
- Pourpanah, F., Abdar, M., Luo, Y., Zhou, X., Wang, R., Lim, C. P., . . . Wu, Q. M. J. (2023). A Review of Generalized Zero-Shot Learning Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4), 4051-4070. <https://doi.org/10.1109/TPAMI.2022.3191696>.
- Presidencia de la República de Costa Rica. (2022). *A dos años de pandemia: Costa Rica se sostiene con fuerza y esperanza gracias al trabajo en equipo de todo el país*. Presidencia de la República de Costa Rica. Recuperado el 30 de abril de 2022, de <https://www.presidencia.go.cr/comunicados/2022/03/a-dos-anos-de-pandemia-costa-rica-se-sostiene-con-fuerza-y-esperanza-gracias-al-trabajo-en-equipo-de-todo-el-pais/>.
- Pretz, K. (31 de marzo de 2021). Stop Calling Everything AI, Machine-Learning Pioneer Says. *IEEE SPECTRUM*. Recuperado el 1 de abril 2021, de <https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says>.
- Princeton University. (2010). *About WordNet*. *WordNet*. Princeton University. Recuperado el 15 de junio de 2022, de <https://wordnet.princeton.edu/>.
- PROCOMER. (3 de febrero de 2021). E-commerce: Las principales tendencias para 2021. Recuperado el 15 de junio de 2022, de [https://www.procomer.com/alertas\\_comerciales/exportador-alerta/e-commerce-las-principales-tendencias-para-2021/](https://www.procomer.com/alertas_comerciales/exportador-alerta/e-commerce-las-principales-tendencias-para-2021/).
- Proyecto Araknion. (2010). *Corpus*. Recuperado el 15 de junio de 2022, de <http://clic.ub.edu/corpus/es/araknion>.
- Pustejovsky, J. y Stubbs, A. (2013). *Natural Language Annotation for Machine Learning* (First ed.). O'Reilly.

- Pustejovsky, J. y Batiukova, O. (2019). *The Lexicon* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. [doi:10.1017/9780511982378](https://doi.org/10.1017/9780511982378).
- Pyysalo, S., Kanerva, J., Missilä, A., Laippala, V. y Ginter, F. (mayo de 2015). Universal dependencies for finnish. In *Proceedings of the 20th Nordic Conference of Computational Linguistics (Nodalida 2015)* (pp. 163-172). Recuperado el 15 de junio de 2022, de <https://aclanthology.org/W15-1821.pdf>.
- Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M. y Yang, D. (2023). Is ChatGPT a general-purpose natural language processing task solver?. *arXiv preprint arXiv:2302.06476*.
- Radford, A., Narasimhan, K., Salimans, T., y Sutskever, I. (2018). Improving language understanding by generative pre-training. Recuperado el 23 de mayo de 2023, de <https://openai.com/research/language-unsupervised>.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., y Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9. Recuperado el 23 de mayo de 2023, de [http://static.cs.brown.edu/courses/cs146/assets/papers/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](http://static.cs.brown.edu/courses/cs146/assets/papers/language_models_are_unsupervised_multitask_learners.pdf).
- Razgovorov, P. y Tomás, D. (2019). Creación de un corpus de noticias de gran tamaño en español para el análisis diacrónico y diatópico del uso del lenguaje. *Procesamiento del Lenguaje Natural*, 62, 29-36. Recuperado el 15 de junio de 2022, de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/5949>.
- Reese, S., Boleda, G., Cuadros, M., Padró, L. y Rigau, G. (mayo de 2010). Wikicorpus: A word-sense disambiguated multilingual wikipedia corpus. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Recuperado el 15 de junio de 2022, de [http://www.lrec-conf.org/proceedings/lrec2010/pdf/222\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2010/pdf/222_Paper.pdf).
- Reiss, M. V. (2023). Testing the Reliability of ChatGPT for Text Annotation and Classification: A Cautionary Remark. *arXiv preprint arXiv:2304.11085*. <https://doi.org/10.48550/arXiv.2304.11085>.
- Research and Development Unit for English Studies, Birmingham City University. (1999). WebCorp: The Web as Corpus. *WebCorp Live Concordance the web in real-time*. Recuperado el 10 de octubre de 2020, de <https://www.webcorp.org.uk/live/>.

- Reyes, José A., Montes, Azucena, González, Juan G. y Pinto, D. E. (2013). Clasificación de roles semánticos usando características sintácticas, semánticas y contextuales. *Computación y Sistemas*, 17(2), 263-272. Recuperado el 02 de septiembre de 2021, de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-55462013000200016&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462013000200016&lng=es&tlng=es).
- Rice, C. (2011). *Massive Modularity, Content Integration, and Language*. *Philosophy of Science*, 78(5), 800-812.
- Robles, V. (2010). *Esquemas de Votación Borda aplicados al etiquetado de Roles Semánticos* [Tesis de maestría]. Universidad de Valencia. Recuperado el 15 de junio de 2022, de <https://riunet.upv.es/handle/10251/14440>.
- Robles, V. (2011). Etiquetado de Roles Semánticos en el marco del corpus CoNLL 09. *Ingenius. Revista de Ciencia y Tecnología*, (5), 63-72. Recuperado el 15 de junio de 2022, de <https://www.redalyc.org/pdf/5055/505554809006.pdf>.
- Rojo, G. (2014). Hispanic Corpus Linguistics. En *The Routledge Handbook of Hispanic Applied Linguistics*. ed. Manel Lacorte, 371–387. Londres y Nueva York: Routledge. Recuperado el 15 de junio de 2022, de <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315882727-32/hispanic-corpus-linguistics-guillermo-rojo>.
- Rojo, G. (2016). Los corpus textuales del español. In J. Gutiérrez-Rexach (Ed.), *Enciclopedia de Lingüística Hispánica* (Vol. 2, pp. 285-296). Routledge. Recuperado el 15 de junio de 2022, de [https://gramatica.usc.es/~grojo/Publicaciones/corpus\\_textuales\\_espanol\\_borrador\\_final.pdf](https://gramatica.usc.es/~grojo/Publicaciones/corpus_textuales_espanol_borrador_final.pdf).
- Rojo, G. (2021). *Introducción a la lingüística de corpus en español*. London: Routledge.
- Rusk, N. (2016). Deep learning. *Nature Methods*, 13(1), 35–35. <https://doi.org/10.1038/nmeth.3707>.
- Sadowsky, S. (2006). Corpus Dinámico del Castellano de Chile (Codicach). Base de datos electrónica. Recuperado el 15 de junio de 2022, de <http://sadowsky.cl/codicach.html>.
- Salguero Lamillar, F.J. (1995). Lógica y análisis del lenguaje natural. En Á. Nepomuceno Fernández (Ed.), *Lógica formal: orígenes, métodos y aplicaciones* (pp. 87-106). Sevilla: Kronos. Recuperado el 15 de junio de 2022, de

<https://idus.us.es/bitstream/handle/11441/70386/Logica%20y%20 analisis%20del%20 lenguaje.pdf?sequence=1>.

- Sánchez-Junquera, J., Rosso, P., Montes-y-Gómez, M. y Chulvi, B. (2021). Masking and BERT-based Models for Stereotype Identification. *Procesamiento Del Lenguaje Natural*, 67, 83-94. Recuperado el 15 de julio de 2022, de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6379>
- Sancho Escrivá, J. V., Fanjul Peyró, C., de la Iglesia Vayá, M., Montell, J. A. y Escartí Fabra, M. J. (2020). Aplicación de la Inteligencia Artificial con Procesamiento del Lenguaje Natural para textos de investigación cualitativa en la relación médico-paciente con enfermedad mental mediante el uso de tecnologías móviles. *Revista de Comunicación y Salud*, 10(1), 19-41. Recuperado el 15 de junio de 2022, de <https://elastic-elbakyan.82-223-83-170.plesk.page/index.php/reys/article/view/199>.
- Sandí Salazar, M., Marín Raventós, G. y Casasola Murillo, E. (2016). Automatización del análisis sintáctico para el español con el fin de crear un treebank estandarizado. *Kañina*, 40(4), 163-174. Recuperado el 15 de junio de 2022, de <https://revistas.ucr.ac.cr/index.php/kanina/article/view/30232>.
- Sharoff, S. (n.d.). Leeds collection of Internet corpora. Querían Internet corpora. Recuperado el 10 de octubre 10 de 2020, de <http://corpus.leeds.ac.uk/internet.html>.
- Shum, H. Y., He, X. D. y Li, D. (2018). From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1), 10-26. <https://doi.org/10.1631/FITEE.1700826>.
- Siegel, D. J. (2007). *La mente en desarrollo: cómo interactúan las relaciones y el cerebro para modelar nuestro ser* (J. Aldekoa, Trans.). Editorial Desclée de Brouwer. Recuperado el 15 de junio de 2022, de [https://www.psicoterapiabilbao.es/wp-content/uploads/2015/11/La\\_mente\\_en\\_desarrollo\\_Siegel.pdf](https://www.psicoterapiabilbao.es/wp-content/uploads/2015/11/La_mente_en_desarrollo_Siegel.pdf).
- Sparkes, M. (2022). Has Google's LaMDA artificial intelligence really achieved sentience? *New Scientist*. Recuperado el 15 de junio de 2022, de <https://www.newscientist.com/article/2323905-has-googles-lamda-artificial-intelligence-really-achieved-sentience/>.
- Spinde, T. (diciembre de 2021). An interdisciplinary approach for the automated detection and visualization of media bias in news articles. In *2021 International Conference on Data*

- Mining Workshops (ICDMW)* (pp. 1096-1103). IEEE. Recuperado el 19 de mayo de 2023, de <https://ieeexplore.ieee.org/document/9679994>.
- Taulé, M., Martí, M. A. y Recasens, M. (2008). AnCora: Multilevel Annotated Corpora for Catalan and Spanish. En *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco. European Language Resources Association (ELRA). Recuperado el 15 de junio de 2022, de [http://www.lrec-conf.org/proceedings/lrec2008/pdf/35\\_paper.pdf](http://www.lrec-conf.org/proceedings/lrec2008/pdf/35_paper.pdf).
- Taulé, M., Peris, A. y Rodríguez, H. (2016). Iarg-AnCora: Spanish corpus annotated with implicit arguments. *Lang Resources y Evaluation* 50, 549–584. <https://doi.org/10.1007/s10579-015-9334-3>.
- Törnberg, P. (2023). Chatgpt-4 outperforms experts and crowd workers in annotating political twitter messages with zero-shot learning. *arXiv preprint arXiv:2304.06588*. <https://doi.org/10.48550/arXiv.2304.06588>.
- Turing, A. M. (1950). I.—Computing Machinery And Intelligence. *Mind*, LIX(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>.
- Turing, A. M. (1956). Can a Machine Think. *The World of Mathematics*, vol. 4, pp. 2099–2123.
- University of Essex. (1998). The Brown Corpus. Recuperado el 9 de octubre de 2020, de: [https://www1.essex.ac.uk/linguistics/external/clmt/w3c/corpus\\_ling/content/corpora/list/private/brown/brown.html](https://www1.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/corpora/list/private/brown/brown.html).
- Üstün, A., Bisazza, A., Bouma, G. y van Noord, G. (2020). UDapter: Language Adaptation for Truly Universal Dependency Parsing. *arXiv preprint arXiv:2004.14327*. <https://doi.org/10.48550/arXiv.2004.14327>.
- Vaamonde, G. (2008). Algunos problemas concretos en la anotación de papeles semánticos. Breve estudio comparativo a partir de los datos de AnCora, SenSem y ADESSE. *Procesamiento del lenguaje natural*, (41), 233-240. Recuperado el 15 de junio de 2022, de <https://www.redalyc.org/pdf/5157/515751741026.pdf>.
- Vázquez, G., Alonso, L., Capilla, J. A., Castellón, I. y Fernández, A. (2006). SenSem: sentidos verbales, semántica oracional y anotación de corpus. *Procesamiento del Lenguaje Natural* (37), 113-119. Recuperado el 15 de junio de 2022, de <https://www.redalyc.org/pdf/5157/515751737015.pdf>.

- Veglis, A. y Maniou, T. A. (2019). Chatbots on the rise: A new narrative in journalism. *Stud. Media Commun*, 7, 1-6. <https://doi.org/10.11114/smc.v7i1.3986>.
- Viber Media S.à r.l. (2010-2021). Rakuten Viber Messenger (16.5.0.18) [Aplicación móvil]. Google Play. Recuperado el 15 de junio de 2022, de [https://play.google.com/store/apps/details?id=com.viber.voip&hl=es\\_CR&gl=US](https://play.google.com/store/apps/details?id=com.viber.voip&hl=es_CR&gl=US).
- Vila, M., Bertran, M., Martí, M. A. y Rodríguez, H. (2015). Corpus annotation with paraphrase types: new annotation scheme and inter-annotator agreement measures. *Language Resources and Evaluation*, 49 (1), 77-105. <https://doi.org/10.1007/s10579-014-9272-5>.
- Völkel, S. y Kretschmar, F. (2021). *Introducing Linguistic Research*. Cambridge: Cambridge University Press. [doi:10.1017/9781316884485](https://doi.org/10.1017/9781316884485).
- Wallis, S. y Nelson, G. (2001). Knowledge discovery in grammatically analysed corpora. *Data Mining and Knowledge Discovery*, 5(4), 305-335. <https://doi.org/10.1023/A:1011453128373>.
- Wang, C., Castellón, I. y Comelles, E. (2020). Linguistic analysis of datasets for semantic textual similarity. *Digital Scholarship in the Humanities*, 35(2), 471-484. <https://doi.org/10.1093/llc/fqy076>.
- Wang, Z., Xie, Q., Ding, Z., Feng, Y., y Xia, R. (2023). Is ChatGPT a good sentiment analyzer? A preliminary study. *arXiv preprint arXiv:2304.04339*. <https://doi.org/10.48550/arXiv.2304.04339>.
- Weizenbaum, J. (1966). ELIZA — A Computer Program for the Study of Natural Language Communication between Man and Machine. *Communications of the Association for Computing Machinery* 9: 36-45. Recuperado el 15 de junio de 2022, de <https://dl.acm.org/doi/pdf/10.1145/365153.365168>.
- Wescoat, M. T. (1987). Practical instructions for working with the formalism of Lexical Functional Grammar. *Course Material for the Lexical Functional Grammar*, 3-39. Recuperado el 15 de junio de 2022, de <https://www1.essex.ac.uk/linguistics/external/LFG/www-lfg.stanford.edu/pubs/papers/wescoat/wescoat-1989-0000.pdf>.
- Whatsapp Inc. (2010-2021). WhatsApp Messenger (2.21.22.26) [Aplicación móvil]. Google Play. [https://play.google.com/store/apps/details?id=com.whatsapp&hl=es\\_CR&gl=US](https://play.google.com/store/apps/details?id=com.whatsapp&hl=es_CR&gl=US).
- Wilcox, E. G., Futrell, R., y Levy, R. (2022). Using computational models to test syntactic learnability. *Linguistic Inquiry*, 1-88. [https://doi.org/10.1162/ling\\_a\\_00491](https://doi.org/10.1162/ling_a_00491).

- Wissler, L., Almashraee, M., Monett, D. y Paschke, A. (2014). The Gold Standard in Corpus Annotation. In *5th IEEE Germany Student Conference*, Passau, Germany. Recuperado el 15 de junio de 2022, de [https://www.researchgate.net/profile/Mohammed-Almashraee/publication/265048097\\_The\\_Gold\\_Standard\\_in\\_Corpus\\_Annotation/links/53fc9d110cf22f21c2f3ef2a/The-Gold-Standard-in-Corpus-Annotation.pdf](https://www.researchgate.net/profile/Mohammed-Almashraee/publication/265048097_The_Gold_Standard_in_Corpus_Annotation/links/53fc9d110cf22f21c2f3ef2a/The-Gold-Standard-in-Corpus-Annotation.pdf).
- World Health Organization. (5 de mayo de 2023). *Declaración Acerca de la decimoquinta Reunión del Comité de Emergencias del Reglamento Sanitario Internacional (2005) sobre la pandemia de enfermedad por coronavirus (COVID-19)*. World Health Organization. Recuperado el 9 de mayo de 2023, de [https://www.who.int/es/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/es/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic).
- Zafra, S., Gómez, J. M. y Navarro-Colorado, B. (2017). Diseño, compilación y anotación de un corpus para la detección de mensajes suicidas en redes sociales. *Procesamiento del Lenguaje Natural*, 59, 65-72. Recuperado el 15 de junio de 2022, de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/5494>.
- Zhang, B., Ding, D. y Jing, L. (2022). How would Stance Detection Techniques Evolve after the Launch of ChatGPT?. *arXiv preprint arXiv:2212.14548*. <https://doi.org/10.48550/arXiv.2212.14548>.
- Zhong, Q., Ding, L., Liu, J., Du, B. y Tao, D. (2023). Can ChatGPT understand too? A comparative study on ChatGPT and fine-tuned BERT. *arXiv preprint arXiv:2302.10198*. <https://doi.org/10.48550/arXiv.2302.10198>.
- Zhou, L., Gao, J., Li, D. y Shum, H. Y. (2020). The design and implementation of xiaoice, an empathetic social chatbot. *Computational Linguistics*, 46(1), 53-93. [https://doi.org/10.1162/coli\\_a\\_00368](https://doi.org/10.1162/coli_a_00368).
- Zhu, Y., Zhang, P., Haq, E. U., Hui, P., & Tyson, G. (2023). Can chatgpt reproduce human-generated labels? a study of social computing tasks. *arXiv preprint arXiv:2304.10145*. <https://doi.org/10.48550/arXiv.2304.10145>.
- Ziems, C., He, B., Soni, S. y Kumar, S. (2020). Racism is a Virus: Anti-Asian Hate and Counterhate in Social Media during the COVID-19 Crisis. *arXiv preprint arXiv:2005.12423*. <https://doi.org/10.48550/arXiv.2005.12423>.

## Anexo A

### Guía de anotación de corpus lingüístico

El objetivo de esta encuesta es etiquetar los argumentos del verbo para cada oración. Para esto, se le entregará un documento EXCEL con tres hojas y cada una de estas hojas contiene una lista de oraciones. Cada oración tendrá un ID, un LEMA y una frase con un segmento subrayado (ver tabla 1). El LEMA es el verbo de la oración que se está analizando, por lo que es necesario tenerlo siempre en cuenta para etiquetar correctamente el argumento que se le pide en cada oración. En algunos casos, el verbo que se estudia coincide con el verbo principal de oración; mientras que, en otros, se trata del verbo de una oración subordinada. En cada ítem, se deberá marcar con una equis (X) el rol temático correspondiente al argumento que se subraya en rojo. Cada oración deberá tener únicamente una opción seleccionada.

**Tabla 1**

*Ejemplos de etiquetado para la anotación del corpus*

ID	LEMA	Frase	agente	tema	meta	experimentador	instrumento	fuelle	lugar	tiempo
3	PERDER	<u>Mi</u> <u>esposa</u> perdió el conocimiento.						X		
4	NACER	<u>En</u> <u>1998</u> , nacieron en Barcelona 11.659 niños.								X

Se etiquetarán los argumentos verbales con ocho posibles opciones, de acuerdo con la función semántica que realice en la oración que se le pida anotar, según tabla 1: el agente es quien realiza la acción; el tema es a quien le ocurre la acción; la meta es el lugar o persona hacia donde se dirige un movimiento o acción; el experimentador es quien percibe o experimenta sensaciones provocadas por una fuente; el instrumento es con qué o con quién se realiza una acción; la fuente provoca sensaciones en otra entidad o es el origen de donde parte una acción; el lugar es donde se desarrolla la acción; y el tiempo es el período relacionado con el desarrollo de la acción.

## Anexo B

### Encuesta para los anotadores

1. \*El infectólogo Carlos Álvarez\* afirmó que ante síntomas gripales lo mejor es aislarse ya que podría ser COVID-19. **(AFIRMAR)**
2. \*El nuevo coronavirus, que ahora se conoce con el nombre de 2019-nCoV o COVID-19\*, no se había detectado antes de que se notificara el brote en Wuhan (China) en diciembre de 2019. **(DETECTAR)**
3. Si \*la persona\* tose y tiene otros síntomas leves, no debe olvidarse de usar una mascarilla al entrar en contacto con otras personas y limpiar las superficies con las que tiene contacto. **(TOSER)**
4. Las dosis de vacuna adicionales (vacunas de refuerzo) pueden proporcionar \*una protección más duradera contra el COVID-19\*. **(PROPORCIONAR)**
5. Si empeoras o tienes dificultad respiratoria o no se controla \*la fiebre\*, llama al 112. **(CONTROLAR)**
6. Para el 1 de enero de 2020, \*el mercado\* había sido cerrado y se había descartado que el causante de la neumonía fuera el SARS, el MERS, gripe, gripe aviaria u otras enfermedades respiratorias comunes causadas por virus. **(CERRAR)**
7. A esto le sumamos que debido al teletrabajo y demanda de trabajos a tiempo parcial ha aumentado y está más solicitado por mujeres que tienen que cuidar \*de sus hijos\* y hacer las tareas domésticas al mismo tiempo. **(CUIDAR)**
8. Un informe publicado por la Organización Mundial de la Salud en febrero descubrió que \*más del 33% de 55.924 personas con casos confirmados por laboratorio de covid-19\* habían tosido esputo, un moco espeso a veces llamado flema, de sus pulmones. **(TOSER)**
9. El Instituto Nacional de Salud (INS) ha dispuesto \*una serie de videos con las preguntas más frecuentes que suele surgir de la mortal enfermedad\*. **(DISPONER)**
10. Downing Street pide disculpas \*a Isabel II\* por las fiestas celebradas en la víspera del funeral de su marido. **(PEDIR)**
11. Este fármaco fue creado \*en la Universidad Federal de los Urales en Ekaterimburgo\* para tratar la gripe H5N1 (gripe aviar). **(CREAR)**
12. El nuevo coronavirus, que ahora se conoce con el nombre de 2019-nCoV o COVID-19, no se había detectado antes de que se notificara el brote en Wuhan (China) \*en diciembre de 2019\*. **(NOTIFICAR)**
13. \*La mayoría de quienes se enferman con COVID-19\* solo presenta una enfermedad leve y puede recuperarse en casa. **(PRESENTAR)**
14. \*El estudio\* describió un subconjunto único de casos más leves en los que los síntomas iniciales fueron problemas digestivos como diarrea, a menudo sin fiebre. **(DESCRIBIR)**
15. Durante una crisis de salud, la desinformación contribuye a que \*la gente\* no se proteja y se haga vulnerable a la enfermedad. **(PROTEGER)**
16. La idea es que las personas eventualmente utilicen \*esta información\* para decidir si necesitan hacerse una prueba de Covid-19 o aislarse. **(UTILIZAR)**
17. Downing Street pide \*disculpas\* a Isabel II por las fiestas celebradas en la víspera del funeral de su marido. **(PEDIR)**
18. Usar pañuelos desechables para eliminar \*secreciones\* y tirarlos después de su uso. **(ELIMINAR)**
19. \*La mayoría de las personas con la enfermedad\* tienen síntomas leves a moderados y se recuperan por completo. **(TENER)**
20. Informó de que había visitado regularmente \*un mercado local de productos frescos\* en Wuhan antes de la aparición de la enfermedad el 5 de enero de 2020. **(VISITAR)**

21. La investigación clínica encontró que se detecta una alta concentración de citoquinas en el plasma de pacientes críticos infectados con SARS-CoV-2, lo que sugiere \*que la tormenta de citoquinas se asociaba con la gravedad de la enfermedad\*. **(SUGERIR)**
22. Sin embargo, con ómicron, \*la mialgia\* podría considerarse uno de los primeros indicios de que los pacientes se habían contagiado con la nueva versión del virus. **(CONSIDERAR)**
23. Una mujer china de 61 años vecina de Wuhan que el 8 de enero viajó en vuelo directo de Wuhan a Tailandia \*en compañía de cinco miembros de su familia\* en un grupo de dieciséis personas. **(VIAJAR)**
24. La tos seca, la fiebre y el cansancio fueron \*los primeros síntomas que\* se asociaron al coronavirus cuando irrumpió en nuestras vidas. **(ASOCIAR)**
25. En España se están administrando \*cuatro vacunas: dos basadas en la tecnología de ARN-m( Pfizer y Moderna) y dos hechas a partir de adenovirus (AstraZeneca y Janssen)\*. **(ADMINISTRAR)**
26. Desecha \*la mascarilla usada\* y lávate las manos. **(DESECHAR)**
27. Si \*la enfermedad\* no se trata a tiempo, estos síntomas pueden agravarse**(TRATAR)**
28. \*Jeremy Konyndyk, exdirector de la Oficina de USAID de Asistencia para Desastres Externos de Estados Unidos\* escribió en un artículo de 2017 que una nueva gran crisis sanitaria global es una cuestión de 'cuándo', no de 'sí', y añadió que el virus sería comparable al de la pandemia de 1918. **(ESCRIBIR)**
29. Si bien la evolución a los casos moderados y graves puede seguir siendo una realidad para cualquiera de las versiones del virus, \*los datos preliminares\* indican que la variante ómicron no causaría enfermedad más grave en los adultos y que la vacuna reduciría aún más la probabilidad de desarrollar cuadros graves, la hospitalización y la muerte, aunque aún hay incertidumbre, informa Sanidad. **(INDICAR)**
30. La bestia sale de noche, dijo Cuomo, refiriéndose a los escalofríos, dolores en el cuerpo y fiebre alta que lo visitaron \*el 1 de abril\*. **(VISITAR)**
31. La mayoría de quienes se enferman con COVID-19 solo presenta \*una enfermedad leve\* y puede recuperarse en casa. **(PRESENTAR)**
32. Para el día de hoy no se han reportado \*defunciones nuevas\*. **(REPORTAR)**
33. Limpie y desinfecte las superficies de contacto frecuente \*en la habitación del enfermo y el baño\*; use guantes desechables. **(DESINFECTAR)**
34. Una persona infectada puede propagar el COVID-19 desde 48 horas (o 2 días) antes de que \*la persona\* haya manifestado algún síntoma o dado positivo en la prueba de detección. **(MANIFESTAR)**
35. Los médicos pueden enviar a los pacientes leves a recuperarse \*en su casa\*, bajo las siguientes condiciones: **(ENVIAR)**
36. El Instituto Feinstein de Northwell Health anunció en marzo \*un estudio sobre un anticuerpo humano que puede prevenir la actividad de IL-6\*. **(ANUNCIAR)**
37. Mientras que los pacientes sin patologías previas tenían una tasa de mortalidad del 1,4 %, \*aquellos que sí tenían\* fallecían a tasas mucho más altas: 13,2 % para enfermedades cardiovasculares, 9,2 % para diabetes, 8,4 % para hipertensión, 8,0 % para enfermedad respiratoria crónica y 7,6 % para cáncer. **(FALLECER)**
38. Después de que los investigadores prueban a fondo \*nuevos tratamientos o vacunas\* en el laboratorio para asegurarse de que puedan beneficiar a las personas, los tratamientos más prometedores pasan a los estudios clínicos. **(PROBAR)**
39. \*Para febrero de 2021\* diez vacunas han sido autorizadas para su uso público por al menos una autoridad reguladora competente. **(AUTORIZAR)**
40. El científico Vaclav Smil escribió en su libro Global Catastrophes and Trends (2008), \*que la posibilidad de una pandemia de influenza durante los próximos 50 años es virtualmente del 100 %\*. **(ESCRIBIR)**
41. \*El virus\* infecta las vías respiratorias y causa síntomas que van desde un cuadro leve (tos seca, fiebre...) a insuficiencia respiratoria aguda y neumonías potencialmente mortales. **(INFECTAR)**

42. De acuerdo con estudios exhaustivos al respecto, sabemos \*que el SRAS-CoV se transmitió de la civeta al ser humano y que se ha producido transmisión del MERS-CoV del dromedario al ser humano\*. **(SABER)**
43. \*La variante delta\*, también conocida como B.1.6172, se ha destacado por su capacidad para replicarse más rápidamente que otras. **(DESTACAR)**
44. Utilizar cubiertas faciales limita \*el volumen y la distancia de viaje de las gotitas que se dispersan al hablar, respirar y toser\*. **(LIMITAR)**
45. Los estudios y los datos actuales muestran que los anticuerpos que \*nuestros cuerpos\* producen después de la vacunación reconocen y protegen contra las variantes del COVID-19. **(PRODUCIR)**
46. Por otro lado, \*la pandemia\* ha afectado a la salud mental, social y sexual de las personas LGTBI por varios motivos. **(AFECTAR)**
47. El ministerio ofreció a los hospitales rusos \*descripciones y guías sobre el mecanismo de acción del tratamiento y las dosis recomendadas\*. **(OFRECER)**
48. \*Los síntomas del resfriado común\* se distinguen fácilmente porque son mucho más leves que los de la gripe y la COVID-19 y comienzan de forma más gradual. **(COMENZAR)**
49. Si tiene COVID-19 y tiene un alto riesgo de enfermedad grave, su médico puede recomendarle \*anticuerpos monoclonales\*. **(RECOMENDAR)**
50. Los casos crecieron rápidamente y el 13 de marzo de 2020 la Organización Mundial de la Salud declaró a Europa el nuevo epicentro del virus después de que \*la situación\* mejorara en China. **(MEJORAR)**
51. Se conoce \*que cualquier persona puede infectarse, independientemente de su edad\*, pero hasta el momento se han registrado relativamente pocos casos de COVID-19 en niños. **(CONOCER)**
52. Entre estos casos de covid-19 atribuidos a la variante ómicron, 34 (79%) ocurrieron en personas que completaron \*la serie primaria de una vacuna contra el covid-19 autorizada o aprobada por la FDA\* 14 días antes del inicio de los síntomas o la recepción de una prueba positiva, incluidos 14 que habían recibido una dosis adicional o de refuerzo. **(COMPLETAR)**
53. En el caso de los adultos, \*el amplio abanico de síntomas, la mayoría con similitudes al de un proceso gripal o catarral\*, puede suponer una mayor dificultad a la hora de determinar el verdadero diagnóstico entre cualquier enfermedad o la COVID. **(SUPONER)**
54. \*La fiebre de la COVID\* suele aparecer más gradualmente. **(APARECER)**
55. Todo el mundo habla \*de la enfermedad por coronavirus 2019 (COVID-19)\*, y allá donde mires hay información sobre el virus y sobre cómo protegerte. **(HABLAR)**
56. \*La investigación\* señala que los problemas oculares más comunes relacionados con la COVID-19 son la sensibilidad a la luz, la irritación o la picazón de los ojos. **(SEÑALAR)**
57. Si su pecho se tensa o comienza a sentir que no puede respirar \*lo suficiente\* como para llenar sus pulmones de aire, esa es una señal para actuar rápidamente, dicen los expertos. **(RESPIRAR)**
58. La idea es que \*las personas\* eventualmente utilicen esta información para decidir si necesitan hacerse una prueba de Covid-19 o aislarse. **(UTILIZAR)**
59. El 31 de diciembre de 2019, la Organización Mundial de la Salud (OMS) recibió \*reportes de presencia de neumonía\*, de origen desconocido, en la ciudad de Wuhan, en China. **(RECIBIR)**
60. Otros pacientes de Wuhan desarrollaron \*dificultad para respirar\* ese día. **(DESARROLLAR)**
61. El SARS-CoV-2 puede sobrevivir y permanecer contagioso \*en superficies inanimadas como metal, vidrio o plástico\* por varios días. **(PERMANECER)**
62. El virus que salió de Wuhan no es exactamente \*el mismo que\* entró en España. **(ENTRAR)**
63. \*Es vital que si tiene alguno de estos síntomas, incluso si está seguro de que podría ser un resfriado, debe reservar inmediatamente y hacerse una prueba de PCR\*, ha reconocido la doctora Philippa Kaye. **(RECONOCER)**
64. El virus que salió \*de Wuhan\* no es exactamente el mismo que entró en España. **(SALIR)**
65. Cúbrase la boca y la nariz \*con un pañuelo desechable o su manga (no sus manos)\* cuando tosa o estornude. **(CUBRIR)**

66. \*Los virus de la gripe y la COVID-19\* provocan enfermedades respiratorias, por lo que algunos de sus síntomas más característicos son similares: fiebre, tos, fatiga, dificultad para respirar, dolor de garganta, mucosidad nasal, dolor muscular y dolor de cabeza. **(PROVOCAR)**
67. Facebook comunicó que eliminaría \*todo contenido que hiciera apología a las protestas contra la inmovilización social\*, el propio Mark Zuckerberg calificó las difusiones de las protestas por redes sociales de información errónea dañina. **(ELIMINAR)**
68. \*En España\* se están administrando cuatro vacunas: dos basadas en la tecnología de ARN-m (Pfizer y Moderna) y dos hechas a partir de adenovirus (AstraZeneca y Janssen). **(ADMINISTRAR)**
69. Asimismo, se ha confirmado que \*otros animales –como perros, gatos, tigres, leones y murciélagos–\* pueden contraer el COVID-19 al igual que los seres humanos. **(CONTRAER)**
70. Por otro lado, los resultados de un estudio publicado en 'The Lancet Microbe', parecen indicar que la mayoría de la transmisión ocurre \*durante los primeros días, especialmente en los 5 primeros después del inicio de los síntomas\*, lo que indica la importancia del confinamiento después del inicio de los síntomas. **(OCURRIR)**
71. Una prueba viral busca \*una infección actual\*. **(BUSCAR)**
72. \*El Sr. Chen\* agregó que si China no notificó a la OMS que era evidencia de un encubrimiento por parte de Beijing, o si China notificó a la OMS, entonces fue un incumplimiento del deber de la organización no transmitir información. **(AGREGAR)**
73. Incluso en el mejor de los casos, desde Naciones Unidas se afirma que las 1300 millones de personas del continente necesitarán \*74 millones de equipos de prueba y 30 000 respiradores\* en 2020. **(NECESITAR)**
74. A nivel inmunológico, se ha evidenciado \*niveles elevados de IL-6 y otras citoquinas proinflamatorias en pacientes con COVID-19 grave\*. **(EVIDENCIAR)**
75. El gobierno federal está proporcionando vacunas contra el COVID-19 autorizadas por la FDA de forma gratuita \*a todas las personas que viven en los Estados Unidos\*, independientemente de su estado migratorio o de seguro médico. **(PROPORCIONAR)**
76. No obstante, actualmente parece ser que es más probable que una persona contagie cuando tiene síntomas, tales como fiebre o tos, pero se desconoce el periodo de tiempo durante el cual \*una persona con SARS\* puede transmitir la enfermedad a otras, es decir, podrían aparecer en tan solo 2 días. **(TRANSMITIR)**
77. Informó de que había visitado regularmente un mercado local de productos frescos en Wuhan antes de la aparición de la enfermedad \*el 5 de enero de 2020\*. **(VISITAR)**
78. No utilice \*el transporte público o taxis\*. **(UTILIZAR)**
79. \*Lo que\* empezó como un quebradero de cabeza para China se ha convertido en un problema de salud pública a nivel mundial, pues aunque la mayoría de casos se registraron originalmente en el país asiático, el epicentro de la pandemia se trasladó primero a Europa, donde los Gobiernos de Italia, España, Alemania y Francia han lidiado con sus respectivas curvas de transmisión; y luego a Estados Unidos, que desde hace semanas es el primer país en número de contagios y muertes registradas. **(EMPEZAR)**
80. La COVID-19 se está propagando rápidamente \*por todo el mundo y dentro de los Estados Unidos\*. **(PROPAGAR)**
81. No reutilice máscaras de un solo uso; deseche \*las máscaras de un solo uso\* después de cada uso y deséchelas inmediatamente después de retirarlas. **(DESECHAR)**
82. Este dato representa \*el porcentaje de personas aptas (5 años o más) que tienen al menos una dosis\*. **(REPRESENTAR)**
83. \*En febrero de 2020\*, Aylward dirigió una delegación de la OMS a Wuhan por la crisis del coronavirus. **(DIRIGIR)**
84. \*Otros pacientes de Wuhan\* desarrollaron dificultad para respirar ese día. **(DESARROLLAR)**
85. Los CDC han actualizado \*algunas de las formas de operar los sistemas de atención médica\* de manera efectiva en respuesta a la vacunación contra el COVID-19. **(ACTUALIZAR)**

86. Cuando el virus entra \*en nuestro cuerpo\* -ya sea por los ojos, la boca o la nariz- se sujeta a las células de la mucosa del fondo de la nariz y la garganta, dice el experto. **(ENTRAR)**
87. \*Los expertos de la OMS\* explican la ciencia relacionada con la COVID-19 en cada episodio. **(EXPLICAR)**
88. El primer ensayo clínico completo, publicado en enero de 2021, concluyó que la aplicación de transfusiones de plasma durante la fase inicial de síntomas leves redujo \*la gravedad de la enfermedad\* en un 60% de los casos. **(REDUCIR)**
89. \*Héctor Raúl Vargas, titular de la División de Medicina Familiar del IMSS\*, explicó que se implementará un código QR que le servirá al paciente saber de inmediato si tiene síntomas de Covid-19. **(EXPLICAR)**
90. El 7 de enero de 2020 los científicos chinos habían aislado el virus causante de la enfermedad, y realizaron \*la secuenciación del genoma\*. **(REALIZAR)**
91. Recientemente se sumaron como síntomas de Covid-19 \*la cefalea, y la presencia de vómitos y/o diarrea\*. **(SUMAR)**
92. El coronavirus causante del covid-19 se ha cobrado la vida de unas 286.000 personas –26.920 de ellas en España– y ha infectado \*a más de 4.177.000\* en todo el mundo, según el último balance global ofrecido por la Universidad Johns Hopkins. **(INFECTAR)**
93. La investigación clínica encontró \*que se detecta una alta concentración de citoquinas\* en el plasma de pacientes críticos infectados con SARS-CoV-2, lo que sugiere que la tormenta de citocinas se asociaba con la gravedad de la enfermedad. **(ENCONTRAR)**
94. El primer ensayo clínico completo, publicado en enero de 2021, concluyó que \*la aplicación de transfusiones de plasma\* durante la fase inicial de síntomas leves redujo la gravedad de la enfermedad en un 60% de los casos. **(REDUCIR)**
95. \*El científico Vaclav Smil\* escribió en su libro Global Catastrophes and Trends (2008), que la posibilidad de una pandemia de influenza durante los próximos 50 años es virtualmente del 100 %. **(ESCRIBIR)**
96. En cuanto a su comienzo, todavía no se ha confirmado \*el posible origen animal de la COVID-19\*. **(CONFIRMAR)**
97. El primer caso de COVID-19 en Europa se detectó el 25 de enero en Francia: \*un hombre de 80 años procedente de la provincia de Hubei, que\* luego falleció el 15 de febrero. **(FALLECER)**
98. Una persona infectada puede propagar el COVID-19 desde 48 horas (o 2 días) antes de que la persona haya manifestado \*algún síntoma\* o dado positivo en la prueba de detección. **(MANIFESTAR)**
99. \*Los pacientes mayores y las personas que tienen afecciones médicas subyacentes o están inmunocomprometidos\* deben comunicarse con su médico temprano incluso en el curso de una enfermedad leve, aconseja el CDC. **(COMUNICAR)**
100. Las declaraciones emitidas por Xi Jinping el 3 de febrero, indicó la necesidad de un énfasis por parte de los medios estatales en contar las historias conmovedoras de cómo \*las personas en la línea del frente\* están previniendo y luchando contra el virus como una prioridad de cobertura, mientras que el alto funcionario Zhang Xiaoguo dijo que su departamento trataría la propaganda con respecto a las medidas de control y prevención del virus como su máxima prioridad. **(PREVENIR)**
101. Aunque el virus ataca primero a los pulmones, también puede infectar \*el corazón, los riñones, el hígado, el cerebro y los intestinos\*. **(INFECTAR)**
102. \*El virus\* comienza un segundo ataque, causando que la condición del paciente empeorar alrededor de 7 a 14 días después del inicio. **(COMENZAR)**
103. Los funcionarios de salud pública deben saber \*cuántas personas están infectadas y dónde viven\* para que los funcionarios puedan rastrear los brotes y tomar medidas para proteger a los demás. **(SABER)**

104. \*Los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) de los Estados Unidos y la Organización Mundial de la Salud (OMS)\* consideran que la COVID-19 es una grave amenaza para la salud pública. **(CONSIDERAR)**
105. El 16 de marzo de 2020, una importante autoridad francesa y asesor del Gobierno francés sobre COVID-19, el profesor Didier Raoult del Instituto Universitario Hospitalario de Enfermedades Infecciosas (IHU-Méditerranée infection) en Marsella (Bouches-du-Rhône, Provenza-Alpes-Côte d'Azur), anunció \*que un ensayo con 24 pacientes del sureste de Francia había demostrado que la cloroquina es un tratamiento efectivo para COVID-19\*. **(ANUNCIAR)**
106. \*El menor de 15 años\* manifestó que sí se quiere vacunar alegando que si se suministra es porque es confiable. **(QUERER)**
107. \*La mayoría de los proveedores de atención médica\* ofrecen pruebas de COVID-19, y debe contactarlos primero a ellos si tiene síntomas. **(OFRECER)**
108. Cualquier persona puede contraer \*la COVID-19\* y caer gravemente enferma. **(CONTRAER)**
109. En Argentina, Canadá, Francia, Alemania, España, Reino Unido y Estados Unidos, las autoridades notificaron \*un incremento de las denuncias de violencia en el hogar\*. **(NOTIFICAR)**
110. Los pacientes mayores y las personas que tienen afecciones médicas subyacentes o están inmunocomprometidos deben comunicarse \*con su médico\* temprano incluso en el curso de una enfermedad leve, aconseja el CDC. **(COMUNICAR)**
111. Si bien la evolución a los casos moderados y graves puede seguir siendo una realidad para cualquiera de las versiones del virus, los datos preliminares indican \*que la variante ómicron no causaría enfermedad más grave en los adultos y que la vacuna reduciría aún más la probabilidad de desarrollar cuadros graves, la hospitalización y la muerte\*, aunque aún hay incertidumbre, informa Sanidad. **(INDICAR)**
112. \*Los análisis preliminares\* también apuntan a que provoca una enfermedad bastante más leve, pero todavía es pronto para extraer conclusiones sobre ómicron. **(APUNTAR)**
113. Agradecemos la iniciativa GISAID, los laboratorios responsables de obtener las muestras, así como \*los laboratorios que\* generaron las secuencias genómicas y compartieron a través de GISAID. **(GENERAR)**
114. Si estás enfermo con la COVID-19, puedes ayudar a prevenir que se propague \*la infección con el virus que causa la COVID-19\*. **(PROPAGAR)**
115. A pesar de las medidas tomadas para frenar la transmisión de contagios y de la campaña de vacunación, desarrollada con un balance muy positivo en España, la variante ómicron del SARS-CoV-2 ha provocado \*una sexta ola con una explosión de casos sin precedentes\*. **(PROVOCAR)**
116. Todo el mundo habla \*del coronavirus\*. **(HABLAR)**
117. Las máscaras limitan el volumen y la distancia de viaje de las gotas respiratorias dispersas \*al hablar, estornudar y toser\*. **(LIMITAR)**
118. Para comparar el orden de los síntomas de la covid-19 con los de la gripe, los investigadores examinaron los datos de 2.470 casos en América del Norte, Europa y el Hemisferio Sur, que se notificaron a las autoridades sanitarias \*entre 1994 y 1998\*. **(NOTIFICAR)**
119. Un informe de la Organización Mundial de la Salud incluye el dolor de garganta entre los síntomas que pueden padecer \*los pacientes con coronavirus\*. **(PADECER)**
120. Así mismo, tanto como el Festival de la Canción de Eurovisión 2020, \*por primera vez en su historia\*, y otros eventos deportivos mundiales como la Eurocopa 2020, la Copa América 2020, la Temporada 2020 de Fórmula 1 y los Juegos Olímpicos de Tokio 2020, han sido cancelados o aplazados hasta este año 2021, entre varios otros eventos mundiales **(CANCELAR)**
121. \*Las dosis de vacuna adicionales (vacunas de refuerzo)\* pueden proporcionar una protección más duradera contra el COVID-19. **(PROPORCIONAR)**
122. Los expertos no ven necesaria \*la cuarta dosis para todos\*. **(VER)**
123. La enfermedad se propaga principalmente de persona a persona a través de \*las gotículas que\* salen despedidas de la nariz o la boca de una persona infectada al toser, estornudar o hablar. **(SALIR)**

124. El gobierno de Hong Kong advirtió a cualquiera que viaje fuera de la ciudad que no toque animales; no coma \*carne de caza\*; y evite visitar mercados húmedos, mercados de aves vivas y granjas. **(COMER)**
125. Hayan pasado al menos 10 días desde que empezaron \*los síntomas\*. **(EMPEZAR)**
126. Rápidamente, a principios de enero, las autoridades de este país identificaron \*la causa como una nueva cepa de coronavirus\*. **(IDENTIFICAR)**
127. \*Los voluntarios adultos que han sido completamente vacunados contra el COVID-19\* reciben dosis de refuerzo de diferentes vacunas contra el COVID-19. **(RECIBIR)**
128. Paralelamente, cabe destacar \*la elevada presencia de mujeres en el ámbito sanitario\*, lo que ha supuesto que aproximadamente el 75% de todo el personal sanitario contagiado en España sean mujeres. **(SUPONER)**
129. \*Las Naciones Unidas\* están siguiendo de cerca la situación del brote de COVID-19. **(SEGUIR)**
130. \*Cualquier persona\* puede contraer la COVID-19 y caer gravemente enferma. **(CAER)**
131. Si no es posible ponerte una mascarilla, cúbrete la boca y la nariz \*con un pañuelo o con el codo\* cuando tosas o estornudes. **(CUBRIR)**
132. En total, se estima que \*haber contraído el virus\* proporciona una inmunización natural contra los síntomas de COVID-19 del 94%. **(PROPORCIONAR)**
133. Si tienes un trastorno de salud mental, como ansiedad o depresión, continúa \*con tu tratamiento\*. **(CONTINUAR)**
134. La sentencia señala \*que la Consejería de Sanidad cometió una vulneración flagrante del principio de igualdad al discriminar al personal sanitario al inicio del plan de vacunación\*. **(SEÑALAR)**
135. Actualmente (agosto 2021), existe \*la recomendación de vacunar a toda mujer embarazada\*. **(EXISTIR)**
136. La revista The Lancet publicó \*el 24 de enero de 2020\* un estudio de los primeros 41 casos de pacientes ingresados (en hospital de Wuhan) con el diagnóstico confirmado, desde el 16 de diciembre de 2019 al 2 de enero de 2020. **(PUBLICAR)**
137. \*Al 3 de marzo de 2020\* a nivel global, 3 110 de los 90 892 casos comunicados de COVID-19 habían muerto. (3,4%), según cifras indicadas por el director de la OMS. **(MORIR)**
138. La persona encargada de los cuidados u otra persona debería usar una mascarilla y guantes desechables \*antes de limpiar\*. **(USAR)**
139. En total, se estima \*que haber contraído el virus proporciona una inmunización natural contra los síntomas de COVID-19 del 94%\*. **(ESTIMAR)**
140. \*Los funcionarios de salud pública\* deben saber cuántas personas están infectadas y dónde viven para que los funcionarios puedan rastrear los brotes y tomar medidas para proteger a los demás. **(SABER)**
141. \*El español\* cree que toda la polémica alrededor de Novak Djokovic no es buena para el tenis. **(CREER)**
142. En otros casos, \*algunas personas infectadas\* no desarrollan ningún síntoma, pero pueden contagiar igualmente al resto de población. **(CONTAGIAR)**
143. La Organización Mundial de la Salud (OMS) apunta que \*cerca de un 80% de los contagiados\* se recupera del covid-19, la enfermedad que provoca el virus, sin necesidad de tratamiento hospitalario. **(RECUPERAR)**
144. Es por ello que los viajes se han visto restringidos y \*muchos países\* han cerrado sus fronteras, para tratar de evitar su expansión. **(CERRAR)**
145. Lo anterior ha ocasionado que los aviones vuelen vacíos entre aeropuertos, se hayan cancelado \*vuelos\*, se hayan cerrado aeropuertos y que la gran mayoría de aviones se encuentren en tierra. **(CANCELAR)**
146. \*Los científicos\* evalúan la seguridad de los regímenes potenciados mixtos y controlan los niveles de anticuerpos resultantes. **(EVALUAR)**
147. Si hay un centro de pruebas y evaluación cerca de ti, puede ir \*allí\* directamente, dijo Schaffer. Siempre es bueno notificarles que irás. **(IR)**

148. Los científicos reconocen sentirse sorprendidos por el nivel y el tiempo que permanecen algunos de los cambios detectables en la sangre que la Covid-19 deja \*en el sistema inmunológico de los afectados\*, incluso en los casos de sintomatología leve. **(DEJAR)**
149. Hay pacientes que son dados de alta, pero \*eso\* no significa que sus síntomas desaparecieron por completo. **(SIGNIFICAR)**
150. Los otros síntomas estén mejorando: la pérdida del gusto y del olfato puede durar \*semanas o meses después de la recuperación\*, pero no es motivo para retrasar el fin del aislamiento. **(DURAR)**
151. Más tarde, uno de los delegados declaró que \*el equipo\* quería llamarlo un patógeno peligroso, y sentían que la palabra peligroso podría estar vinculada al bioterrorismo. **(QUERER)**
152. El coronavirus suele propagarse a través de pequeñas gotas de agua que produce \*una persona con el virus\* cuando habla, tose o estornuda. **(ESTORNUDAR)**
153. Hay que estar alerta si siente \*una dificultad para respirar\*. **(SENTIR)**
154. En Argentina, Canadá, Francia, Alemania, España, Reino Unido y Estados Unidos, \*las autoridades\* notificaron un incremento de las denuncias de violencia en el hogar. **(NOTIFICAR)**
155. La investigación clínica encontró que se detecta una alta concentración de citoquinas en el plasma de pacientes críticos infectados con SARS-CoV-2, \*lo que\* sugiere que la tormenta de citocinas se asociaba con la gravedad de la enfermedad. **(SUGERIR)**
156. \*El estudio, que\* observó los casos de más de 181 pacientes, concluyó que el tiempo medio de incubación es de 5 días, y que los síntomas pueden, por lo menos a largarse, otros 12 días. **(OBSERVAR)**
157. Esta afección, además, es una señal de que \*tus pulmones\* podrían estar sufriendo severamente la infección con el SARS-CoV-2. **(SUFRIR)**
158. \*La anosmia\*, en particular, se ha visto en pacientes que finalmente dieron positivo para el coronavirus sin otros síntomas, según la Academia Americana de Otorrinolaringología-Cirugía de Cabeza y Cuello. **(VER)**
159. Esto causa irritación y por ello empezamos a toser, señala \*Schaffner\*. **(SEÑALAR)**
160. Las personas mayores y las personas con ciertas afecciones de salud existentes tienen \*un mayor riesgo de desarrollar enfermedad grave y morir\*. **(TENER)**
161. Impedir un pico agudo de infecciones, lo que se conoce como achatar la curva de epidemia, reduce \*el riesgo de colapso de los sistemas de salud\* y proporciona más tiempo para el desarrollo de un tratamiento o de una vacuna. **(REDUCIR)**
162. Limpie \*todas las áreas de contacto frecuente\* en el hogar, como las manijas de las puertas, accesorios de baño y cocina, inodoros, teléfonos, tabletas, encimeras y otras superficies. **(LIMPIAR)**
163. Una enfermedad, incluida la COVID-19, no \*siempre\* responde a los mismos síntomas. **(RESPONDER)**
164. Personas que llevan máscaras quirúrgicas \*en Guangzhou\*. **(LLEVAR)**
165. Se acordó con las cámaras empresariales que los empleadores no pedirán \*pruebas positivas o negativas\* para ausentarse de las labores. **(PEDIR)**
166. Los datos analizados por la app mostraron que la tos era el síntoma más común al comienzo de la pandemia, y \*el 46 por ciento de los pacientes infectados\* padecían este signo revelador. **(PADECER)**
167. La DSMB está compuesta por expertos en la afección específica del estudio que periódicamente analizan \*los resultados del estudio\* a medida que avanza. **(ANALIZAR)**
168. Por el momento, no se recomienda ningún medicamento específico para prevenir o tratar la infección, pero \*los médicos de medio mundo\* trabajan a contrarreloj en desarrollar una cura tomando como referencia las experiencias pasadas. **(TRABAJAR)**
169. Lo anterior ha ocasionado que los aviones vuelen vacíos entre aeropuertos, se hayan cancelado vuelos, se hayan cerrado \*aeropuertos\* y que la gran mayoría de aviones se encuentren en tierra. **(CERRAR)**

170. Y aunque los datos mejoran \*en el tercer y cuarto trimestre\*, sigue sin haber una disminución considerable en la violencia hacia las mujeres. **(MEJORAR)**
171. En febrero de 2020, Aylward dirigió una delegación de la OMS \*a Wuhan\* por la crisis del coronavirus. **(DIRIGIR)**
172. Al igual que a la de otros virus, la respuesta inmunitaria a la COVID-19 dependerá \*de si se dispone de un sistema inmunitario sano\*. **(DEPENDER)**
173. Después de que los investigadores prueban a fondo nuevos tratamientos o vacunas en el laboratorio para asegurarse de que puedan beneficiar a las personas, los tratamientos más prometedores pasan \*a los estudios clínicos\*. **(PASAR)**
174. No obstante, actualmente parece ser que es más probable que una persona contagie cuando tiene síntomas, tales como fiebre o tos, pero se desconoce el periodo de tiempo durante el cual una persona con SARS puede transmitir \*la enfermedad\* a otras, es decir, podrían aparecer en tan solo 2 días. **(TRANSMITIR)**
175. Y esto ha llevado a que \*muchos gobiernos\* impongan nuevas restricciones. **(IMPONER)**
176. Y \*los CDC y la OMS\* incluyen la diarrea entre los síntomas más comunes del COVID-19. **(INCLUIR)**
177. Recientemente (diciembre 2021), se ha publicado una revisión sistemática en la que estudios observacionales en los que participaron casi dos millones de adultos sugieren que \*la deficiencia / insuficiencia de vitamina D\* aumenta la susceptibilidad al COVID-19 y al COVID-19 grave. **(AUMENTAR)**
178. Buena parte de las mujeres trabajan \*en el sector de la restauración y del turismo\*, de los más afectados por la pandemia, además de ser las principales cuidadoras de personas enfermas, de niños y de personas de la tercera edad. **(TRABAJAR)**
179. A pesar de que las autoridades del Servicio Nacional de Salud inglés (NHS) reconocen a la temperatura alta, tos persistente y a la pérdida o cambio repentino en el sentido de olfato o gusto. (anosmia y/o disgeusia) como principales síntomas del coronavirus, investigadores del King's College de Londres que dirigen el proyecto de vigilancia más ambicioso del mundo denominado COVID ZOE Symptom Study alertaron que la enfermedad pandémica está \*ahora\* actuando de manera diferente. **(ACTUAR)**
180. A comienzos de 2020, el mundo dirigió la mirada hacia China, donde el SARS-CoV-2 avanzaba \*por todo el país\*. **(AVANZAR)**
181. En tanto, \*quienes deseen hacerse el examen de forma voluntaria y están afiliados o afiliadas a Fonasa\*, podrán realizárselo en un hospital público. **(REALIZAR)**
182. Se acordó con las cámaras empresariales que \*los empleadores\* no pedirán pruebas positivas o negativas para ausentarse de las labores. **(PEDIR)**
183. Esta página compara \*el COVID-19\* con la influenza, con base en la mejor información disponible a la fecha. **(COMPARAR)**
184. \*La Resolución 380 de 2020\* establece que tanto los viajeros nacionales como extranjeros provenientes de China, Italia, Francia y España deben adoptar medidas de aislamiento e internación en su residencia o en un hospedaje transitorio cubierto por su propia cuenta durante 14 días. **(ESTABLECER)**
185. En esta página se recoge \*documentación de la Comunidad de Madrid, el Estado, el Ayuntamiento de Madrid y la Unión Europea\*. **(RECOGER)**
186. \*La sentencia\* señala que la Consejería de Sanidad cometió una vulneración flagrante del principio de igualdad al discriminar al personal sanitario al inicio del plan de vacunación. **(SEÑALAR)**
187. \*A finales de enero\*, el Ministerio de Salud de Rusia identificó tres medicamentos para adultos que podrían ayudar a tratar la enfermedad. **(IDENTIFICAR)**
188. Necesitamos solidarizarnos con las historias entrañables que \*nos\* cuentan en las redes, pero en la inmensidad de internet también necesitamos un recordatorio de que esa historia nos importó un día. **(CONTAR)**

189. Ante la incertidumbre generada, \*la OMS\* ha desmentido que se propague a grandes distancias a través del aire y ha asegurado que la principal vía de contagio son las gotículas que se generan cuando una persona infectada tose o estornuda, o a través de saliva o de secreciones de la nariz. **(ASEGURAR)**
190. Así avanza \*la vacunación\* en España: cuatro de cada diez niños ya tienen la primera dosis de la vacuna. **(AVANZAR)**
191. Los pacientes que pierden el olfato no siempre pierden \*el gusto\*. **(PERDER)**
192. En otros casos, el coronavirus provocó infartos y accidentes cerebrovasculares (ACV) \*en personas jóvenes\* e insensibilidad a los efectos de la falta de oxígeno, que se conoce como hipoxia feliz. **(PROVOCAR)**
193. Este país no lleva \*un conteo de personas recuperadas\*. **(LLEVAR)**
194. De 84 personas a las que se les diagnosticó Covid \*durante el estudio\*, el algoritmo detectó 67. **(DIAGNOSTICAR)**
195. Lo importante es que las personas, incluso si sufren de alergias, resfriados o gripe común, mantengan una distancia de un metro de los demás \*al toser o estornudar\*. **(MANTENER)**
196. Sin embargo, la ausencia de anticuerpos en la sangre no significa que el sistema inmunitario no pueda producir \*nuevos anticuerpos\* rápidamente en caso de reexposición al SARS-CoV-2. **(PRODUCIR)**
197. \*Ya\* hemos hablado de China. **(HABLAR)**
198. Asegúrate de que \*en casa\* todos conocen cómo hacer un correcto lavado de manos. **(CONOCER)**
199. En un futuro se espera \*que también puedan avisar si tienes gripe o la Covid-19 antes de presentar síntomas\*, algo que respalda una nueva investigación. **(ESPERAR)**
200. Saber que cada enfermedad progresa de una manera diferente significa que los médicos pueden identificar antes si es probable que alguien tenga la covid-19 u otro tipo de enfermedad, lo que puede ayudar a tomar mejores decisiones sobre los tratamientos que deban administrarse o puede ayudar a los doctores a descartar \*otras enfermedades\*. **(DESCARTAR)**
201. Investigadores de la Universidad Noruega de Ciencia y Tecnología (NTNU) han creado \*una base de datos con 120 agentes antivirales de amplio espectro seguros para las personas\* e identificaron 31 candidatos a fármacos para el tratamiento de SARS-CoV-2. **(CREAR)**
202. No obstante, actualmente parece ser que es más probable que \*una persona\* contagie cuando tiene síntomas, tales como fiebre o tos, pero se desconoce el periodo de tiempo durante el cual una persona con SARS puede transmitir la enfermedad a otras, es decir, podrían aparecer en tan solo 2 días. **(CONTAGIAR)**
203. \*Investigadores chinos\* descubrieron que Arbidol, un medicamento antiviral utilizado para tratar la gripe, podría combinarse con Darunavir, un medicamento empleado en el tratamiento del VIH, para el tratamiento de pacientes con coronavirus. **(DESCUBRIR)**
204. \*La Consejería de Sanidad\* considerará este resultado verdadero positivo y se recomendará autoaislamiento durante 7 días, evitar recibir visitas y extremar las medidas de prevención habituales. **(CONSIDERAR)**
205. Los estudios están analizando si \*los tratamientos\* acortan el tiempo de recuperación y disminuyen los problemas respiratorios y de otro tipo de la infección por COVID-19. **(DISMINUIR)**
206. Así lo hace \*cada comunidad\*. **(HACER)**
207. Asegúrate de que en casa \*todos\* conocen cómo hacer un correcto lavado de manos. **(CONOCER)**
208. En febrero, más de 140 médicos de cabecera escribieron \*al director médico de Inglaterra Asesor Científico Jefe del Departamento de Salud y Asistencia Social, Chris Whitty\*, instando al gobierno a agregar más síntomas parecidos al resfriado a la lista oficial. **(ESCRIBIR)**
209. No evaluamos \*a personas sin síntomas\* porque tenemos un problema de recursos, dijo Schaffner sobre el centro de evaluación en Vanderbilt. **(EVALUAR)**
210. Según el Centro Chino para el Control y Prevención de Enfermedades (CCDC), el 29 de diciembre un hospital en Wuhan (Hospital Provincial de Medicina Integrada Tradicional China y Occidental,

- también conocido como el hospital de Xinhua), admitió a 4 individuos con neumonía, quienes trabajaban \*en un mercado de esa ciudad\*. **(TRABAJAR)**
211. En cuanto a estado de vacunación, se muestra el número de personas aptas en Puerto Rico que han completado \*una serie de vacunas COVID-19\*. **(COMPLETAR)**
212. \*Quienes ya se estaban recuperando\* pueden ver que sus fiebres desaparecen después de 12 días. **(VER)**
213. Hasta el momento, \*las pruebas que se han realizado\* descartan la presencia de SARS-CoV-2 en el líquido amniótico, en la sangre de cordón umbilical y en la leche materna. **(DESCARTAR)**
214. La reducción de linfocitos B puede ocurrir \*temprano\* en la enfermedad, que puede afectar la producción de anticuerpos en el paciente. **(OCURRIR)**
215. CCINSHAE y Sedesa llamaron a poner atención en seguir adelante con \*las medidas de seguridad que\* eviten la propagación de contagios. **(EVITAR)**
216. Todas las personas son susceptibles de adquirir el virus, siendo las que tienen más riesgo \*aquellas que\* tratan con personas contagiadas. **(TRATAR)**
217. No obstante, actualmente parece ser que es más probable que una persona contagie \*cuando tiene síntomas, tales como fiebre o tos\*, pero se desconoce el periodo de tiempo durante el cual una persona con SARS puede transmitir la enfermedad a otras, es decir, podrían aparecer en tan solo 2 días. **(CONTAGIAR)**
218. Queremos \*conocer\* para poder ofrecerte los contenidos según tus preferencias. **(QUERER)**
219. También estudiaron \*un conjunto de datos de casi 1.100 casos recogidos desde el 11 de diciembre de 2019 hasta el 29 de enero de 2020 por el Grupo de Expertos en Tratamiento Médico de China a través de la Comisión Nacional de Salud de China\*. **(ESTUDIAR)**
220. A pesar de ello, \*los síntomas del contagio de la COVID-19\* siguen manifestándose, cada vez de formas más variadas. **(MANIFESTAR)**
221. La infección se produce cuando \*una persona enferma\* tose o estornuda y expulsa partículas del virus que entran en contacto con otras personas. **(ESTORNUDAR)**
222. Se procuró que \*la denominación\* no contuviera nombres de personas o referencias a ningún lugar, especie animal, tipo de comida, industria, cultura o grupo de personas, en línea con las recomendaciones internacionales, para evitar que hubiera estigmatización contra algún colectivo. **(CONTENER)**
223. De acuerdo con ABC, \*los síntomas del coronavirus\* han cambiado. **(CAMBIAR)**
224. Además, las circunstancias actuales dificultan el acceso a teléfonos, líneas de atención y servicios públicos como la policía, la justicia y los servicios sociales, por lo que elevan \*la vulnerabilidad de las mujeres a la violencia en el ámbito privado en todo el mundo\*. **(ELEVAR)**
225. Al igual que a la de otros virus, \*la respuesta inmunitaria a la COVID-19\* dependerá de si se dispone de un sistema inmunitario sano. **(DEPENDER)**
226. Tras el primer brote de COVID-19 en Wuhan en diciembre de 2019, donde \*las autoridades chinas\* confirmaron 41 casos detectados entre el 8 de diciembre y el 2 de enero de 2020, la ciudad dejó de informar casos hasta el 19 de enero, cuando se confirmaron 17 casos más. **(CONFIRMAR)**
227. \*Las citoquinas son proteínas solubles que\* actúan sobre las células del sistema inmune y que regulan la activación, proliferación y reclutamiento celular. **(ACTUAR)**
228. Para el 1 de enero de 2020, el mercado había sido cerrado y se había descartado \*que el causante de la neumonía fuera el SARS, el MERS, gripe, gripe aviaria u otras enfermedades respiratorias comunes causadas por virus\*. **(DESCARTAR)**
229. Enjuagar la nariz, hacer gárgaras con enjuague bucal y comer \*ajo\* no son métodos efectivos. **(COMER)**
230. En febrero, \*más de 140 médicos de cabecera\* escribieron al director médico de Inglaterra Asesor Científico Jefe del Departamento de Salud y Asistencia Social, Chris Whitty, instando al gobierno a agregar más síntomas parecidos al resfriado a la lista oficial. **(ESCRIBIR)**

231. Tras una última fase de maduración, en la que intervienen proteasas virales, todos los componentes del virus encajarán, la partícula será infecciosa y podrá comenzar \*un nuevo ciclo celular\*. **(COMENZAR)**
232. La arquitecta Carmen Romero propuso llenar el Oso y el Madroño de estrellas, en recuerdo de \*los que\* nos dejaron a causa de la covid**(DEJAR)**
233. \*Este\* le dará las guías locales más recientes acerca del examen. **(DAR)**
234. Además de la guía mencionada anteriormente sobre el lavado de manos y la higiene respiratoria, \*los organismos de salud pública\* aconsejan que las personas enfermas que sospechan que pueden tener COVID-19 deben restringir las actividades fuera del hogar, excepto para obtener atención médica. **(ACONSEJAR)**
235. En general, el coronavirus afecta menos a la parte alta del aparato respiratorio, a diferencia del resfriado común aunque, como decimos, las nuevas variantes han cambiado \*este patrón\*. **(CAMBIAR)**
236. Las medidas en la economía han afectado especialmente a este colectivo debido a que un elevado número de personas LGTBI trabajan en sectores en los que \*las restricciones de contención del virus\* han supuesto la pérdida de numerosos puestos de trabajo, y dependen en muchos casos de fuentes de ingresos informales. **(SUPONER)**
237. La OMS recuerda que las mascarillas sólo son eficaces si se combinan con el lavado frecuente de manos con una solución hidroalcohólica con agua y jabón y únicamente recomienda su uso en personas sanas si \*estas\* atienden a alguien en quien se sospeche la infección o si se sufre tos y estornudos frecuentes. **(ATENDER)**
238. La enfermedad es mortal en raras ocasiones, y hasta ahora las víctimas mortales han sido \*personas de edad avanzada que\* ya padecían una enfermedad crónica como diabetes, asma o hipertensión. **(PADECER)**
239. Los estudios respaldados por los NIH tienen como objetivo determinar si \*los medicamentos llamados anticoagulantes\* pueden tratar a los adultos diagnosticados con COVID-19 y prevenir la formación de coágulos sanguíneos potencialmente mortales. **(PREVENIR)**
240. En Estados Unidos, ya se está probando otra posible vacuna \*en humanos\*. **(PROBAR)**
241. Las personas mayores y \*las que\* padecen afecciones médicas previas como hipertensión arterial, problemas cardíacos o pulmonares, diabetes o cáncer tienen más probabilidades de presentar cuadros graves. **(PADECER)**
242. Entre estos casos de covid-19 atribuidos a la variante ómicron, 34 (79%) ocurrieron en \*personas que\* completaron la serie primaria de una vacuna contra el covid-19 autorizada o aprobada por la FDA 14 días antes del inicio de los síntomas o la recepción de una prueba positiva, incluidos 14 que habían recibido una dosis adicional o de refuerzo. **(COMPLETAR)**
243. \*Durante más de dos meses\*, los españoles sufrieron uno de los confinamientos más estrictos, lo que no evitó que los servicios sanitarios se vieran superados en muchos lugares. **(SUFRIR)**
244. \*La variante ómicron\* se ha notificado hasta ahora en 19 países y territorios de la región. **(NOTIFICAR)**
245. Hay otros signos del coronavirus que no presentan los enfermos de gripe, como la pérdida del gusto o el olfato, las erupciones cutáneas, las diarreas, los cambios de color \*en los dedos de las manos o los pies\*. **(PRESENTAR)**
246. Hay \*otros signos del coronavirus que\* no presentan los enfermos de gripe, como la pérdida del gusto o el olfato, las erupciones cutáneas, las diarreas, los cambios de color en los dedos de las manos o los pies. **(PRESENTAR)**
247. El síndrome de liberación de citoquinas se produce \*cuando se activan grandes cantidades de leucocitos (neutrófilos, macrófagos y mastocitos) y liberan grandes cantidades de citoquinas proinflamatorias\*. **(PRODUCIR)**
248. Es por eso que el Instituto Nacional de Salud (INS) dio recomendaciones de cómo actuar ante la aparición de \*cualquier indicio que\* pueda generar la enfermedad. **(GENERAR)**

249. Por favor, no agregues \*datos especulativos\* y recuerda colocar referencias a fuentes fiables para dar más detalles. **(AGREGAR)**
250. Según el Centro Chino para el Control y Prevención de Enfermedades (CCDC), el 29 de diciembre un hospital en Wuhan (Hospital Provincial de Medicina Integrada Tradicional China y Occidental, también conocido como el hospital de Xinhua), admitió a \*4 individuos con neumonía, quienes\* trabajaban en un mercado de esa ciudad. **(TRABAJAR)**
251. La RT-PCR da positivo durante varias semanas después de la primera infección (30 días de media, según algunos estudios), ya que detecta la presencia del ARN del virus, aunque este ya no sea viable y el paciente haya superado \*la infección\* y ya no sea contagioso. **(SUPERAR)**
252. Las máscaras también se recomiendan para \*quienes\* cuidan a alguien que pueda tener la enfermedad. **(CUIDAR)**
253. \*La mayoría de quienes se enferman con COVID-19\* solo presenta una enfermedad leve y puede recuperarse en casa. **(RECUPERAR)**
254. En el caso de los adultos, el amplio abanico de síntomas, la mayoría con similitudes al de un proceso gripal o catarral, puede suponer \*una mayor dificultad a la hora de determinar el verdadero diagnóstico entre cualquier enfermedad o la COVID\*. **(SUPONER)**
255. A medida que se propagaba la pandemia, se cancelaban o posponían \*las conferencias y eventos mundiales en tecnología, moda, deportes, entre otros\*. **(CANCELAR)**
256. \*Esto\* conllevó a terminar una fase de preparación en el territorio nacional y empezar la fase de contención que era desarrollada por el Comité Nacional de Emergencias. **(EMPEZAR)**
257. El tocilizumab (un anticuerpo monoclonal) ha sido incluido en las pautas de tratamiento por la Comisión Nacional de Salud de China después de que se completó \*un pequeño estudio\*. **(COMPLETAR)**
258. El paracetamol (Tylenol) y el ibuprofeno. (Advil, Motrin) ayudan \*a reducir la fiebre\*. **(AYUDAR)**
259. La variante ómicron del COVID-19 se expande rápidamente por el mundo y el Perú no es la excepción, \*los estudios\* apuntan a que es la más contagiosa hasta la fecha. **(APUNTAR)**
260. \*Una persona acatarrada\* tiene congestión nasal, estornudos y puede que fiebre baja y tos. **(TENER)**
261. También es esencial que usen su antebrazo o un pañuelo. (que deben desechar) cuando tosen o estornudan y que tras hacerlo \*siempre\* se laven las manos para evitar la propagación de otros virus, advierte. **(LAVAR)**
262. Por el contrario, \*los síntomas más comunes en jóvenes vacunados son fiebre, dolor de cabeza, tos persistente y dolor de garganta\*, reconoció el profesor Duncan. **(RECONOCER)**
263. La medida más importante de prevención es mantener la distancia social para frenar \*la transmisión del virus\*. **(FRENAR)**
264. Algunas investigaciones han sugerido \*que COVID-19 es una enfermedad vascular en lugar de respiratoria\*, lo que significa que puede viajar a través de los vasos sanguíneos. **(SUGERIR)**
265. Conforme a los hallazgos de la investigación británica, \*1 de cada 5 pacientes\* experimenta también dolores en la zona abdominal. **(EXPERIMENTAR)**
266. Puedes tener también \*pérdida de olfato\* y más adelante falta de aire. **(TENER)**
267. La novedad surge \*del rastreo que hace la aplicación COVID ZOE Symptom Study, el mayor monitoreo comunitario del mundo\*. **(SURGIR)**
268. Es por ello que los viajes se han visto restringidos y muchos países han cerrado \*sus fronteras\*, para tratar de evitar su expansión. **(CERRAR)**
269. Si \*sus síntomas\* no indican una emergencia médica, su proveedor puede decidir controlar sus síntomas mientras se recupera en casa. **(INDICAR)**
270. Suelen ser síntomas leves y comienzan gradualmente, por lo que en un principio es la fiebre y la tos \*lo primero que\* experimenta el infectado de COVID-19. **(EXPERIMENTAR)**

271. \*Impedir un pico agudo de infecciones, lo que se conoce como achatar la curva de epidemia\*, reduce el riesgo de colapso de los sistemas de salud y proporciona más tiempo para el desarrollo de un tratamiento o de una vacuna. **(REDUCIR)**
272. Esto incluye comer \*en tu cuarto\*. **(COMER)**
273. \*La llegada del nuevo coronavirus\* a América Latina ha generado dudas sobre la diferencia entre los síntomas de la enfermedad covid-19 y los de una gripe común. **(GENERAR)**
274. Desarrollada con el apoyo de los NIH, \*esta vacuna\* no requiere refrigeración especial. **(REQUERIR)**
275. \*Los pacientes que están hospitalizados con una enfermedad grave por COVID-19\* también pueden recibir anticoagulantes para prevenir o tratar los coágulos de sangre. **(RECIBIR)**
276. Por lo tanto, \*muchas mujeres quienes ya sufrían unas condiciones precarias\* perderán su empleo. **(PERDER)**
277. \*Especialistas\* apuntan a que esta tendencia se debe a que nuestra sociedad sigue asociando a las mujeres a las labores de cuidados, aunque sea de forma inconsciente. **(APUNTAR)**
278. Separarse de otras personas y animales \*en el hogar\*; no compartir artículos personales; usar un baño separado si está disponible. **(SEPARAR)**
279. Si estás cuidando a alguien con COVID-19 y no estás completamente vacunado, los CDC recomiendan que hagas cuarentena durante 14 días después del último contacto \*con la persona enferma\* y que estés atento a cualquier síntoma de la COVID-19. **(RECOMENDAR)**
280. Puedes tener también pérdida de olfato y \*más adelante\* falta de aire. **(TENER)**
281. Se reportaron casos de reinfección en pacientes que ya habían pasado \*la enfermedad\* y se temía que el SARS-CoV-2 se comportase como algunos otros coronavirus, que son capaces de reinfectar después de aproximadamente un año. **(PASAR)**
282. Así que la recomendación del experto es que si se tienen síntomas gripales se aisle inmediatamente y llame \*a la EPS\* para solicitar una prueba de COVID-19. **(LLAMAR)**
283. A veces, \*las personas que están completamente vacunadas\* contraen una infección irruptiva, lo que significa que dan positivo en la prueba del SARS-CoV-2 o se enferman con COVID-19. **(CONTRAER)**
284. \*En febrero\*, más de 140 médicos de cabecera escribieron al director médico de Inglaterra Asesor Científico Jefe del Departamento de Salud y Asistencia Social, Chris Whitty, instando al gobierno a agregar más síntomas parecidos al resfriado a la lista oficial. **(ESCRIBIR)**
285. Si le diagnostican \*COVID-19\*, es posible que alguien del departamento de salud lo llame. **(DIAGNOSTICAR)**
286. La revista The Lancet publicó el 24 de enero de 2020 \*un estudio de los primeros 41 casos de pacientes ingresados (en hospital de Wuhan) con el diagnóstico confirmado, desde el 16 de diciembre de 2019 al 2 de enero de 2020\*. **(PUBLICAR)**
287. Esto conllevó a terminar una fase de preparación en el territorio nacional y empezar \*la fase de contención que era desarrollada por el Comité Nacional de Emergencias\*. **(EMPEZAR)**
288. Los estudios respaldados por los NIH tienen como objetivo determinar si los medicamentos llamados anticoagulantes pueden tratar a los adultos diagnosticados con COVID-19 y prevenir \*la formación de coágulos sanguíneos potencialmente mortales\*. **(PREVENIR)**
289. \*El 2 de febrero de 2020\*, médicos de Tailandia afirmaron haber tratado a un paciente con éxito con una combinación de lopinavir / ritonavir y el medicamento contra la influenza oseltamivir. **(AFIRMAR)**
290. La OMS anunció el 11 de febrero de 2020 \*que COVID-19 sería el nombre oficial de la enfermedad\*. **(ANUNCIAR)**
291. Así que los vacunados todavía deben realizar resoluciones preventivas (lavarse las manos con jabón, ponerse mascarilla, mantener la distancia social, etc.) para que \*la enfermedad\* no se propague, especialmente hacia gente vulnerable. **(PROPAGAR)**
292. El problema no es solo la infección, sino la forma en que responde \*nuestro organismo\* para luchar contra ella, le explica a BBC Mundo Kalpana Sabapathy médica clínica y epidemióloga del equipo

- de salud global de la Escuela de Higiene y Medicina Tropical en Londres, Reino Unido. **(RESPONDER)**
293. Este artículo trata \*sobre la descripción médica de la enfermedad\*. **(TRATAR)**
294. A veces \*el primer caso de COVID-19 identificado por el sistema sanitario\* (caso índice o paciente 0), no coincide con el caso que da origen a la epidemia. **(COINCIDIR)**
295. \*Las personas mayores y las personas con ciertas afecciones de salud existentes\* tienen un mayor riesgo de desarrollar enfermedad grave y morir. **(TENER)**
296. Creo que el mensaje aquí es que si eres joven y tienes cualquier síntoma leve, aunque pueda parecer un mal resfriado o una sensación rara... quédate \*en casa\* y hazte una prueba, señala el profesor Spector. **(QUEDAR)**
297. Asimismo, se debe evitar \*el contacto estrecho\* con cualquier persona que presente signos de afección respiratoria, como tos o estornudos. **(EVITAR)**
298. Para el caso de pacientes con sintomatología grave, se les referirá \*a un hospital del IMSS\* para ser atendidos a la brevedad, expuso Vargas Sánchez. **(REFERIR)**
299. \*La gente\* no se está dando cuenta de este cambio y podría pensar que tiene algún tipo de resfriado estacional, van a fiestas y pueden contagiar. **(PENSAR)**
300. \*Aquí\* encontrará información actualizada. **(ENCONTRAR)**
301. Desarrollada con el apoyo de los NIH, esta vacuna no requiere \*refrigeración especial\*. **(REQUERIR)**
302. \*ONU\* destaca la labor de la Organización en 2021 para hacer frente al COVID-19, incluido el programa respaldado por la ONU para ayudar a los países en desarrollo a proteger a sus poblaciones contra el virus, y a las medidas adoptadas para prepararse para futuras crisis sanitarias mundiales. **(DESTACAR)**
303. Necesitamos solidarizarnos con \*las historias entrañables que\* nos cuentan en las redes, pero en la inmensidad de internet también necesitamos un recordatorio de que esa historia nos importó un día. **(CONTAR)**
304. \*Este mapa\* dispone de un buscador (recuadro arriba a la izquierda) para localizar una calle o domicilio; el resultado sitúa la zona que corresponde. **(DISPONER)**
305. La transmisión del SARS-CoV-2 se produce mediante pequeñas gotas –microgotas de Flügge– que se emiten al hablar, estornudar, toser o espirar, que al ser despedidas por un portador (que puede no tener síntomas de la enfermedad o estar incubándola) pasan directamente a otra persona mediante la inhalación, o quedan sobre los objetos y superficies que rodean al emisor, y luego, a través de las manos, que lo recogen \*del ambiente contaminado\*, toman contacto con las membranas mucosas orales, nasales y oculares, al tocarse la boca, la nariz o los ojos. **(RECOGER)**
306. Comparte \*esta información\* con tu familia, tus amigos y tus colegas para conseguir que todo el mundo conozca todos los datos relativos al COVID-19 y puedan proteger su salud. **(COMPARTIR)**
307. La OMS desaconsejaba en marzo la utilización de máscara quirúrgica por la población sana, en abril la OMS consideró \*que era una medida aceptable en algunos países\*. **(CONSIDERAR)**
308. Hacerse pruebas también proporciona \*información importante sobre el movimiento del virus dentro y entre comunidades\*. **(PROPORCIONAR)**
309. Un estudio de más de 550 hospitales en China también identificó que los pacientes hospitalizados desarrollaron \*neumonía\* al tercer día de su enfermedad. **(DESARROLLAR)**
310. Además, este tipo de medidas han agravado \*la situación de soledad en miembros de la comunidad, particularmente en personas mayores LGTBI\*, ya que frecuentemente viven solas y no tienen relaciones positivas con sus familias. **(AGRAVAR)**
311. El nuevo Coronavirus causa \*una Infección Respiratoria Aguda (IRA)\*, es decir una gripa, que puede ser leve, moderada o severa. **(CAUSAR)**
312. Este tipo de epidemias suelen estar provocadas por virus de animales que usan ácido ribonucleico. (ARN) como material genético para mutar y acaban \*en otros seres vivos\*. **(ACABAR)**

313. El Nuevo Coronavirus COVID-19 es una cepa de la familia de coronavirus que no se había identificado previamente \*en humanos\*. **(IDENTIFICAR)**
314. Los fabricantes de vacunas envían actualizaciones de seguridad mensuales \*a la FDA\*. **(ENVIAR)**
315. Tenían una ventaja porque \*ya\* sabían mucho sobre cómo funcionan las vacunas de ARNm y sobre cómo fabricarlas. **(SABER)**
316. Mayo Clinic es una organización sin fines de lucro, y el dinero recaudado con la publicidad en Internet apoya \*nuestra misión\*. **(APOYAR)**
317. Si dudas y no \*te\* han hecho el test, mejor extrema el aislamiento. **(HACER)**
318. \*Este dato\* representa el porcentaje de personas aptas (5 años o más) que tienen al menos una dosis. **(REPRESENTAR)**
319. Saber que cada enfermedad progresa de una manera diferente significa que los médicos pueden identificar antes si es probable que alguien tenga la covid-19 u otro tipo de enfermedad, lo que puede ayudar a tomar mejores decisiones sobre \*los tratamientos que\* deban administrarse o puede ayudar a los doctores a descartar otras enfermedades. **(ADMINISTRAR)**
320. Esto significa \*que la enfermedad se está extendiendo en varios países de todo el mundo\*, al mismo tiempo afectando a cientos de miles de personas en todos los continentes, causando un número de muertes que, a la fecha, sigue en ascenso. **(SIGNIFICAR)**
321. Un estudio de Wuhan descubrió \*que, en promedio, pasaron 12 días hasta que los pacientes fueron admitidos en la UCI\*. **(DESCUBRIR)**
322. Los Centros para el Control y la Prevención de Enfermedades (CDC) no pueden dar fe sobre la precisión de un sitio web que no pertenece \*al Gobierno Federal\*. **(PERTENECER)**
323. El ministerio ofreció \*a los hospitales rusos\* descripciones y guías sobre el mecanismo de acción del tratamiento y las dosis recomendadas. **(OFRECER)**
324. Todos quienes viven \*con el paciente\* deben ser monitoreados por la Autoridad Sanitaria. **(VIVIR)**
325. Si estás cuidando a alguien con COVID-19 y no estás completamente vacunado, los CDC recomiendan que hagas cuarentena \*durante 14 días después del último contacto\* con la persona enferma y que estés atento a cualquier síntoma de la COVID-19. **(RECOMENDAR)**
326. Los expertos de la OMS explican la ciencia relacionada con la COVID-19 \*en cada episodio\*. **(EXPLICAR)**
327. Incluso en el mejor de los casos, desde Naciones Unidas se afirma que las 1300 millones de personas del continente necesitarán 74 millones de equipos de prueba y 30 000 respiradores \*en 2020\*. **(NECESITAR)**
328. \*El aislamiento\* se utiliza para separar a las personas que tienen el virus de la COVID-19 de las personas que no están enfermas. **(UTILIZAR)**
329. Las infecciones asintomáticas son las más comunes entre niños, pero las afecciones dependen de la edad, ya que el dolor estómago es un síntoma común en los más pequeños, mientras que \*los mayores de 12 años\* pueden sufrir una pérdida del gusto o del olfato, según reconocieron un grupo de expertos en epidemiología. **(SUFRIR)**
330. Para comparar el orden de los síntomas de la covid-19 con los de la gripe, los investigadores examinaron \*los datos de 2.470 casos en América del Norte, Europa y el Hemisferio Sur, que\* se notificaron a las autoridades sanitarias entre 1994 y 1998. **(NOTIFICAR)**
331. Al igual que para las salas de cine, se prevé que muchos espacios de exhibición no puedan superar \*un largo periodo de inactividad\*. **(SUPERAR)**
332. Otros tratamientos que se consideraron prometedores durante los primeros meses de la pandemia, como la hidroxiclороquina y el lopinavir/ritonavir, resultaron \*después\* ser inefectivos o incluso perjudiciales. **(RESULTAR)**
333. Que los animales de compañía propaguen el nuevo coronavirus ha sido también negado por la institución, que no tiene ninguna prueba de que \*el nuevo coronavirus\* infecte a animales de compañía como los perros y los gatos. **(INFECTAR)**
334. Las células B de memoria específicas para las proteínas de la nucleocápside y el pico del SARS-CoV-2 duran \*al menos 6 meses después de la aparición de los síntomas\*. **(DURAR)**

335. Durante más de dos meses, \*los españoles\* sufrieron uno de los confinamientos más estrictos, lo que no evitó que los servicios sanitarios se vieran superados en muchos lugares. **(SUFRRIR)**
336. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron \*los casos\* al CCDC, que envió expertos a Wuhan para apoyar la investigación. **(COMUNICAR)**
337. \*El 18 de marzo\* un artículo informa que el tratamiento con lopinavir/ritonavir da negativo en pruebas clínicas con 199 pacientes en China. **(INFORMAR)**
338. \*En un futuro\* se espera que también puedan avisar si tienes gripe o la Covid-19 antes de presentar síntomas, algo que respalda una nueva investigación. **(ESPERAR)**
339. Es por eso que el Instituto Nacional de Salud (INS) dio recomendaciones de cómo actuar ante la aparición de cualquier indicio que pueda generar \*la enfermedad\*. **(GENERAR)**
340. \*El Supremo de EE.UU.\*, señala que Biden no puede obligar a los empleados de grandes empresas a vacunarse. **(SEÑALAR)**
341. De igual modo, \*en Sudán del Sur\*, la situación se agravó considerablemente durante el confinamiento. **(AGRAVAR)**
342. Al estornudar, cúbrete \*con la parte interna del codo\*. **(CUBRIR)**
343. Se está investigando \*el uso de donaciones de plasma sanguíneo de personas que ya se han recuperado de la COVID-19\*, una estrategia que se aplicó con éxito contra la gripe de 1918 y que se ha probado para el SARS, un primo anterior de COVID-19 y otras enfermedades como la fiebre hemorrágica argentina. **(INVESTIGAR)**
344. Puedes tener también pérdida de olfato y más adelante \*falta de aire\*. **(TENER)**
345. La mayoría de las personas tienen una enfermedad leve y logran \*recuperarse\* en casa. **(LOGRAR)**
346. \*Esta variante\* parece funcionar de manera ligeramente diferente, explica el epidemiólogo. **(FUNCIONAR)**
347. Dado que el SARS-CoV-2 ha estado en la población humana solo desde diciembre de 2019, \*al principio de la pandemia\* se desconocía si la inmunidad era duradera en las personas que se recuperan de la enfermedad. **(DESCONOCER)**
348. Esto significa que la enfermedad se está extendiendo en varios países de todo el mundo, al mismo tiempo afectando a cientos de miles de personas en todos los continentes, causando \*un número de muertes que\*, a la fecha, sigue en ascenso. **(SEGUIR)**
349. Las medidas en la economía han afectado especialmente a este colectivo debido a que un elevado número de personas LGTBI trabajan en sectores en los que las restricciones de contención del virus han supuesto \*la pérdida de numerosos puestos de trabajo\*, y dependen en muchos casos de fuentes de ingresos informales. **(SUPONER)**
350. Italia mantiene la alerta \*en apertura de tráfico aéreo 3:33\*. **(MANTENER)**
351. Desde el día 2 de marzo se incluye información sobre la evolución de la vacunación; los datos de vacunación no se actualizan \*los fines de semana y festivos\*. **(ACTUALIZAR)**
352. Teníamos que encontrar un nombre que no se refiriera \*a una ubicación geográfica, un animal, un individuo o a un grupo de personas\*, y que también fuera pronunciable y relacionado con la enfermedad, explicó Adhanom. **(REFERIR)**
353. También es frecuente \*el dolor de cabeza, la fiebre, el olor reducido, la respiración fuerte, el dolor abdominal y el dolor muscular\*, aunque, por lo general, suelen desaparecer pasados dos días. **(DESAPARECER)**
354. Deseche el pañuelo \*después de usarlo\*. **(DESECHAR)**
355. Con él coincide \*Adolfo García Sastre, que dirige en Nueva York el Instituto Global de Salud y Patógenos Emergentes en la Escuela de Medicina de Icahn en el Hospital Mount Sinai\*. **(COINCIDIR)**
356. \*(HHS)\* está desarrollando un plan para distribuir dosis adicionales de vacunas de ARNm a quienes califiquen para recibirlas. **(DESARROLLAR)**

357. Los pacientes que están hospitalizados con una enfermedad grave por COVID-19 también pueden recibir \*anticoagulantes\* para prevenir o tratar los coágulos de sangre. **(RECIBIR)**
358. Los médicos pueden determinar \*qué medidas tomar para cuidar al paciente\*, y pueden evitar que la condición del paciente empeore, resalta uno de los líderes de la investigación, Peter Kuhn. **(DETERMINAR)**
359. Las máscaras quirúrgicas (normales) son utilizadas desde hace tiempo \*por personas sanas\* en Hong Kong, Japón, Singapur y Malasia. **(UTILIZAR)**
360. Se conoce que cualquier persona puede infectarse, independientemente de su edad, pero \*hasta el momento\* se han registrado relativamente pocos casos de COVID-19 en niños. **(REGISTRAR)**
361. Debido a la complejidad y coste de las pruebas PCR, en los primeros meses de la pandemia casi ningún país dispuso \*de datos confiables sobre la prevalencia del virus en su población\*. **(DISPONER)**
362. \*En Italia\* se cuentan como víctimas de la pandemia a los muertos que habían dado positivo pero no a muchos ancianos fallecidos en residencias sin haberles hecho la prueba; mientras que en Bélgica se contabiliza en las cifras de fallecidos a todos los muertos en residencias con síntomas similares a los de COVID-19, aunque no se les hayan hecho tests. **(CONTAR)**
363. La publicación recuerda \*que el Servicio Nacional de Salud (NHS) británico actualmente solo menciona tres síntomas oficiales de covid: la fiebre, tos y cambios en el gusto o el olfato\*. **(RECORDAR)**
364. Los investigadores de los NIH también están probando \*regímenes de vacunas mixtas\*. **(PROBAR)**
365. El test debe realizarse antes de acudir a la reunión, dado que \*el resultado de la prueba\* solo tarda 15 minutos. **(TARDAR)**
366. El Gobierno vasco impone \*a la productora musical Last Tour\* la máxima sanción por no exigir el pasaporte covid en los recitales y permitir que nadie llevara protección. **(IMPONER)**
367. La investigación señala \*que los problemas oculares más comunes relacionados con la COVID-19 son la sensibilidad a la luz, la irritación o la picazón de los ojos\*. **(SEÑALAR)**
368. Informó de que había visitado regularmente un mercado local de productos frescos en Wuhan \*antes de la aparición de la enfermedad\* el 5 de enero de 2020. **(VISITAR)**
369. Limpia \*tu casa\* frecuentemente. **(LIMPIAR)**
370. \*Recientemente\* se sumaron como síntomas de Covid-19 la cefalea, y la presencia de vómitos y/o diarrea. **(SUMAR)**
371. Hasta el momento, las pruebas que se han realizado descartan \*la presencia de SARS-CoV-2 en el líquido amniótico, en la sangre de cordón umbilical y en la leche materna\*. **(DESCARTAR)**
372. El 7 de enero de 2020 \*los científicos chinos\* habían aislado el virus causante de la enfermedad, y realizaron la secuenciación del genoma. **(REALIZAR)**
373. Así son los síntomas del coronavirus día a día, \*un desglose\* muestra la evolución de cada uno. **(MOSTRAR)**
374. Las personas mayores y las que padecen \*afecciones médicas previas como hipertensión arterial, problemas cardíacos o pulmonares, diabetes o cáncer\* tienen más probabilidades de presentar cuadros graves. **(PADECER)**
375. \*Para ese entonces\* ya se habían comunicado los primeros casos por COVID-19 fuera de China: dos en Tailandia y uno en Japón. **(COMUNICAR)**
376. El 18 de diciembre de 2020, la FDA autorizó \*el uso de emergencia de esta vacuna contra el COVID-19 financiada por los NIH en los Estados Unidos para personas mayores de 18 años\*. **(AUTORIZAR)**
377. \*Este fármaco\* fue creado en la Universidad Federal de los Urales en Ekaterimburgo para tratar la gripe H5N1 (gripe aviar). **(CREAR)**
378. Un estudio de más de 550 hospitales en China también identificó que los pacientes hospitalizados desarrollaron neumonía \*al tercer día de su enfermedad\*. **(DESARROLLAR)**

379. Con esta información se puede obtener gráficos para comparar de manera visual \*el modo en que se ha visto impactada la movilidad a raíz de la pandemia\*. **(COMPARAR)**
380. En promedio, las personas en Wuhan se recuperaron del virus y fueron dadas de alta del hospital o fallecieron \*después de 2,5 a 3 semanas\*. **(FALLECER)**
381. El Oso y Madroño de la artista Noelia Jiménez, teñido de una amplia paleta de colores vivos, representa \*cómo los trabajadores esenciales han unido a la sociedad española durante la pandemia\***(REPRESENTAR)**
382. \*La partícula viral\* saldrá de la célula por exocitosis. **(SALIR)**
383. Los otros síntomas estén mejorando: \*la pérdida del gusto y del olfato\* puede durar semanas o meses después de la recuperación, pero no es motivo para retrasar el fin del aislamiento. **(DURAR)**
384. \*El virus SARS-CoV-2\* puede ingresar al cuerpo cuando otra persona inhala estas gotas. **(INGRESAR)**
385. \*Otros estudios\* probarán si los medicamentos utilizados para tratar otras afecciones podrían ayudar a tratar el COVID-19 en personas con síntomas leves a moderados. **(PROBAR)**
386. \*En una encuesta\*, más del 50% de las mujeres con responsabilidades familiares respondieron que se encargaban ellas solas de la enseñanza de sus hijos o de ayudar a algún familiar con dependencia, mientras que en el caso de los hombres solo el 8% afirmaban que en su casa ocurriera este hecho. **(RESPONDER)**
387. Sin embargo, Michael Snyder, uno de los autores del estudio, aseguró que el hallazgo más importante fue que pudieron detectar \*el 80% de los casos confirmados de coronavirus\* antes o cuando los participantes presentaban síntomas. **(DETECTAR)**
388. \*Estos fármacos\* se usan habitualmente para tratar la hepatitis C, infección por VIH y la esclerosis múltiple, respectivamente. **(USAR)**
389. La pandemia de COVID-19 en América es una epidemia continental que inició con la detección del primer caso de esta enfermedad el 21 de enero de 2020 en los Estados Unidos, un hombre de aproximadamente 30 años, originario del estado de Washington, que había viajado recientemente \*a China\*. **(VIAJAR)**
390. Cuando tú o tu ser querido se recuperen, busca \*apoyo emocional\*. **(BUSCAR)**
391. El profesor Tim Spector, epidemiólogo del King's College de Londres y que dirige \*el llamado Estudio de Síntomas Zoe Covid\*, explica que las personas más jóvenes que contraen la variante delta pueden sentirse más como si tuvieran un resfriado fuerte. **(DIRIGIR)**
392. Los investigadores han probado el nuevo algoritmo, desarrollado con software de código abierto, en \*2.155 usuarios que\* durante ocho meses llevaron relojes inteligentes capaces de medir el estrés y las alteraciones en la frecuencia cardíaca, que si es más alta de lo normal puede ser señal clara de que existe algún problema. **(LLEVAR)**
393. A veces, \*las personas que están completamente vacunadas\* contraen una infección irruptiva, lo que significa que dan positivo en la prueba del SARS-CoV-2 o se enferman con COVID-19. **(ENFERMAR)**
394. El primer ensayo clínico completo, publicado en enero de 2021, concluyó que la aplicación de transfusiones de plasma \*durante la fase inicial de síntomas leves\* redujo la gravedad de la enfermedad en un 60% de los casos. **(REDUCIR)**
395. Para comparar el orden de los síntomas de la covid-19 con los de la gripe, los investigadores examinaron los datos de 2.470 casos en América del Norte, Europa y el Hemisferio Sur, que se notificaron \*a las autoridades sanitarias\* entre 1994 y 1998. **(NOTIFICAR)**
396. Si la DSMB encuentra que \*el tratamiento experimental\* no está funcionando o está perjudicando a los participantes, la Junta recomendará que el estudio se detenga de inmediato. **(FUNCIONAR)**
397. La infección se produce cuando una persona enferma tose o estornuda y expulsa \*partículas del virus que\* entran en contacto con otras personas. **(ENTRAR)**
398. Específicamente, 182 países continuaban \*la interrupción total de los servicios de educación\* y 8 lo hacían parcialmente de modo local en distintas jurisdicciones. **(CONTINUAR)**

399. \*A finales de marzo\*, Estados Unidos se convirtió en el país más afectado del mundo. **(CONVERTIR)**
400. Use una mascarilla de tela con al menos 2 capas cuando visite \*el consultorio o la sala de emergencia\*, a menos que le resulte muy difícil respirar. **(VISITAR)**
401. Todo el mundo habla de la enfermedad por coronavirus 2019 (COVID-19), y \*allá donde mires\* hay información sobre el virus y sobre cómo protegerte. **(HABER)**
402. Pudimos determinar que los pacientes que no habían recibido Plaquenil (el medicamento que contiene \*hidroxicloroquina\*) seguían siendo contagiosos después de seis días, pero de los que habían recibido Plaquenil, después de seis días, solo el 25% seguía siendo contagioso. **(CONTENER)**
403. COVID-19 es una enfermedad altamente infecciosa que suele producir síntomas que una persona puede controlar \*en casa\*. **(CONTROLAR)**
404. Así, por ejemplo, la iniciativa Global 50/50 recoge \*en su página web\* información de 55 países, de los cuales únicamente 26 desglosan por sexo los datos actualizados de casos confirmados y de muertes. **(RECOGER)**
405. Todos debemos colaborar para frenar \*la transmisión de la enfermedad por coronavirus\*. **(FREJAR)**
406. Los estudios también indican que \*la protección contra el SARS-CoV-2\* comienza a disminuir con el tiempo después de las dosis iniciales de vacuna. **(DISMINUIR)**
407. Hay casos confirmados de coronavirus \*en casi todo del mundo, incluyendo China, Europa (siendo en Italia y España los primeros focos), Irán, Estados Unidos, África, Australia y LATAM\*. **(HABER)**
408. Si tú o la persona que tiene COVID-19 presentan signos que advierten que se trata \*de una emergencia\*, se necesita atención médica inmediata. **(TRATAR)**
409. \*El ministerio\* ofreció a los hospitales rusos descripciones y guías sobre el mecanismo de acción del tratamiento y las dosis recomendadas. **(OFRECER)**
410. \*El 16 de marzo de 2020\*, una importante autoridad francesa y asesor del Gobierno francés sobre COVID-19, el profesor Didier Raoult del Instituto Universitario Hospitalario de Enfermedades Infecciosas (IHU-Méditerranée infection) en Marsella (Bouches-du-Rhône, Provenza-Alpes-Côte d'Azur), anunció que un ensayo con 24 pacientes del sureste de Francia había demostrado que la cloroquina es un tratamiento efectivo para COVID-19. **(ANUNCIAR)**
411. Respecto a la aparición de los síntomas tras el contagio, la OMS destaca \*que la media es de 5 o 6 días\*, aunque dependiendo de la cepa pueden aparecer hasta los 14 días. **(DESTACAR)**
412. El primer caso de COVID-19 en Europa se detectó el 25 de enero en Francia: un hombre de 80 años procedente de la provincia de Hubei, que \*luego\* falleció el 15 de febrero. **(FALLECER)**
413. Teníamos que encontrar \*un nombre que\* no se refiriera a una ubicación geográfica, un animal, un individuo o a un grupo de personas, y que también fuera pronunciable y relacionado con la enfermedad, explicó Adhanom. **(REFERIR)**
414. \*Utilizar cubiertas faciales\* limita el volumen y la distancia de viaje de las gotitas que se dispersan al hablar, respirar y toser. **(LIMITAR)**
415. \*Un baño tibio o un baño de esponja\* puede ayudar a bajar la fiebre. **(AYUDAR)**
416. Afirma que es imposible poner en un balanza \*los efectos adversos\* a medio y largo plazo. **(PONER)**
417. Asegúrate de que \*en casa\* todos conocen cómo hacer un correcto lavado de manos. **(ASEGURAR)**
418. En otros casos, el coronavirus provocó \*infartos y accidentes cerebrovasculares (ACV) en personas jóvenes e insensibilidad a los efectos de la falta de oxígeno, que se conoce como hipoxia feliz\*. **(PROVOCAR)**
419. En promedio, \*las personas en Wuhan\* se recuperaron del virus y fueron dadas de alta del hospital o fallecieron después de 2,5 a 3 semanas. **(FALLECER)**

420. Resultados de investigación mostraron que la proteína ORF8 del SARS-CoV-2 y la glicoproteína de superficie podrían unirse a la porfirina, respectivamente, mientras que las proteínas del SARS-CoV-2 orf1ab, ORF10 y ORF3a podrían atacar de forma coordinada el hemo para disociar el hierro para formar \*la porfirina\*. **(FORMAR)**
421. La mayoría de quienes se enferman con COVID-19 solo presenta una enfermedad leve y puede recuperarse \*en casa\*. **(RECUPERAR)**
422. En Estados Unidos, \*ya\* se está probando otra posible vacuna en humanos. **(PROBAR)**
423. Limpie \*el área u objeto\* con agua y jabón u otro detergente si está sucio. **(LIMPIAR)**
424. Y aunque \*esas consecuencias\* no resulten muy dañinas, lo cierto es que pueden ser muy duraderas, ya que los investigadores hablan, incluso, de la existencia de secuelas incluso diez semanas después de que el virus SARS-CoV-2 haya desaparecido del organismo del paciente. **(RESULTAR)**
425. Sin embargo, la Food and Drug Administration (FDA) recomienda que \*las mujeres embarazadas o lactantes\* consulten con sus respectivos médicos a la hora de decidir si deben vacunarse. **(CONSULTAR)**
426. Para el caso de pacientes con sintomatología grave, se \*les\* referirá a un hospital del IMSS para ser atendidos a la brevedad, expuso Vargas Sánchez. **(REFERIR)**
427. Los estudios muestran \*que las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19\*, incluso para personas con alto riesgo de contraer la enfermedad. **(MOSTRAR)**
428. Esta afección dermatológica también fue \*uno de los primeros síntomas extraños\* que se identificaron tras el contagio con el SARS-CoV-2. **(IDENTIFICAR)**
429. Lavarse las manos con frecuencia \*durante 20 segundos\* con agua y jabón es fundamental para el cuidador. **(LAVAR)**
430. \*El síndrome de liberación de citoquinas\* se produce cuando se activan grandes cantidades de leucocitos (neutrófilos, macrófagos y mastocitos) y liberan grandes cantidades de citoquinas proinflamatorias. **(PRODUCIR)**
431. Según un estudio noruego publicado en la revista Eurosurveillance, los síntomas más frecuentes son \*la tos, la congestión nasal y la fatiga, que\* duran de media cuatro días. **(DURAR)**
432. A comienzos de 2020, el mundo dirigió la mirada \*hacia China\*, donde el SARS-CoV-2 avanzaba por todo el país. **(DIRIGIR)**
433. En este estudio se están evaluando \*las diferentes dosis de esta vacuna experimental\* para la seguridad y su capacidad de inducir una respuesta inmune en los voluntarios. **(EVALUAR)**
434. Este tipo de epidemias suelen estar provocadas por \*virus de animales que\* usan ácido ribonucleico. (ARN) como material genético para mutar y acaban en otros seres vivos. **(ACABAR)**
435. \*El 31 de diciembre de 2019\*, la Organización Mundial de la Salud (OMS) recibió reportes de presencia de neumonía, de origen desconocido, en la ciudad de Wuhan, en China. **(RECIBIR)**
436. A pesar de las medidas tomadas para frenar la transmisión de contagios y de la campaña de vacunación, desarrollada con un balance muy positivo en España, \*la variante ómicron del SARS-CoV-2\* ha provocado una sexta ola con una explosión de casos sin precedentes. **(PROVOCAR)**
437. Un informe de la Organización Mundial de la Salud incluye \*el dolor de garganta\* entre los síntomas que pueden padecer los pacientes con coronavirus. **(INCLUIR)**
438. Si \*le\* diagnostican COVID-19, es posible que alguien del departamento de salud lo llame. **(DIAGNOSTICAR)**
439. \*El virus de Wuhan\* se parece mucho al SARS, que no tuvo un claro comportamiento estacional. **(PARECER)**
440. Esta inflamación pulmonar aumentada estaría asociada \*al síndrome de dificultad respiratoria aguda (SDRA)\* que se ha descrito como la principal causa de mortalidad por COVID-19. **(DESCRIBIR)**
441. Comparte esta información con tu familia, tus amigos y tus colegas para conseguir que \*todo el mundo conozca\* todos los datos relativos al COVID-19 y puedan proteger su salud. **(CONOCER)**

442. Ir al médico en caso de fiebre, tos y dificultad para respirar, llamando con antelación si se encuentra en zonas donde se está propagando el virus o si se las han visitado \*en los últimos 14 días\*. **(VISITAR)**
443. Saber cómo actuar nos puede ayudar a controlar mejor las situaciones que puedan surgir \*estos días\* y a prestar ayuda a las personas de nuestro entorno. **(SURGIR)**
444. Otros pacientes podrían haber sufrido esta afección, pero no se incluye entre \*las afecciones que\* la Organización Mundial de la Salud (OMS) asocia al COVID-19. **(ASOCIAR)**
445. Así, por ejemplo, la iniciativa Global 50/50 recoge en su página web \*información de 55 países\*, de los cuales únicamente 26 desglosan por sexo los datos actualizados de casos confirmados y de muertes. **(RECOGER)**
446. Según estudios que se han hecho sobre el SARS (síndrome respiratorio agudo grave), primo del covid-19, pensamos \*que puede ir a otras partes del cuerpo\*, dice Schaffner. **(PENSAR)**
447. Si no dispones de agua ni jabón, usa un desinfectante de manos que contenga por lo menos \*un 60% de alcohol\*. **(CONTENER)**
448. \*España\* es el segundo por número de contagios, aunque ya empieza a ver la luz al final del túnel. **(VER)**
449. \*En abril\*, se anunció que la ivermectina inhibe la replicación del SARS-CoV-2 in vitro. **(ANUNCIAR)**
450. De hecho, \*muchos síntomas de covid-19\* pueden parecerse a la gripe, incluidos los dolores de cabeza y los problemas digestivos, dolores corporales y fatiga mencionados anteriormente. **(PARECER)**
451. Si los síntomas parecen estar empeorando, llama \*al médico\*. **(LLAMAR)**
452. Nos esforzamos por lograr \*transparencia y precisión\* en nuestros datos. **(LOGRAR)**
453. Las personas también pueden aplicar \*métodos de distanciamiento social\* limitando los viajes, evitando áreas llenas de gente y alejándose físicamente de las personas enfermas. **(APLICAR)**
454. \*Estos efectos secundarios raros\* generalmente aparecen dentro de las seis semanas después de recibir la dosis de la vacuna. **(APARECER)**
455. Los científicos reconocen sentirse sorprendidos por el nivel y el tiempo que permanecen algunos de los cambios detectables en la sangre que \*la Covid-19\* deja en el sistema inmunológico de los afectados, incluso en los casos de sintomatología leve. **(DEJAR)**
456. Sin embargo, ahora se están realizando estudios clínicos para probar la seguridad y eficacia de las vacunas disponibles actualmente \*en niños más pequeños\*. **(REALIZAR)**
457. Una posible señal de que podrías tener covid-19 es si tus síntomas no mejoran \*después de una semana\* más o menos, pero en realidad empeoran. **(MEJORAR)**
458. Con el cierre de la actividad regular algunas productoras han cedido material a los hospitales y \*gran número de creadores, actores y bailarines\* han acudido al mundo digital, en especial a las redes sociales, para respaldar las diversas iniciativas sociales, reivindicar el sector o simplemente aportar nuevo contenido cultural y seguir haciendo arte. **(ACUDIR)**
459. La DSMB está compuesta por expertos en la afección específica del estudio que \*periódicamente\* analizan los resultados del estudio a medida que avanza. **(ANALIZAR)**
460. \*Una enfermedad\*, incluida la COVID-19, no siempre responde a los mismos síntomas. **(RESPONDER)**
461. Lo que compartimos \*en el ardor del momento\* depende de nosotros. **(COMPARTIR)**
462. El tiempo hace la diferencia, por ello es necesario que revises tus síntomas y acudas \*al hospital\* de acuerdo con el semáforo de atención. **(ACUDIR)**
463. Es vital que si tiene alguno de estos síntomas, incluso si está seguro de que podría ser un resfriado, debe reservar inmediatamente y hacerse una prueba de PCR, ha reconocido \*la doctora Philippa Kaye\*. **(RECONOCER)**
464. Hay \*otros signos del coronavirus que no presentan los enfermos de gripe\*, como la pérdida del gusto o el olfato, las erupciones cutáneas, las diarreas, los cambios de color en los dedos de las manos o los pies. **(HABER)**

465. El Supremo de EE.UU. señala \*que Biden no puede obligar a los empleados de grandes empresas a vacunarse\*(**SEÑALAR**)
466. \*El Instituto Nacional de Salud (INS)\* ha dispuesto una serie de videos con las preguntas más frecuentes que suele surgir de la mortal enfermedad. (**DISPONER**)
467. Los estudios realizados en Reino Unido, de seguimiento de síntomas desde los inicios de la pandemia, evidenciaron durante la ola de delta que \*los dolores de cabeza\* se habían convertido en el síntoma más común de la infección con la variante. (**CONVERTIR**)
468. Los estudios están analizando \*si los tratamientos acortan el tiempo de recuperación y disminuyen los problemas respiratorios y de otro tipo de la infección por COVID-19\*. (**ANALIZAR**)
469. \*Para el 1 de enero de 2020\*, el mercado había sido cerrado y se había descartado que el causante de la neumonía fuera el SARS, el MERS, gripe, gripe aviaria u otras enfermedades respiratorias comunes causadas por virus. (**DESCARTAR**)
470. \*Este país\* no lleva un conteo de personas recuperadas. (**LLEVAR**)
471. Otros estudios probarán \*si los medicamentos utilizados para tratar otras afecciones podrían ayudar a tratar el COVID-19 en personas con síntomas leves a moderados\*. (**PROBAR**)
472. \*Cualquier persona\* puede contraer la COVID-19 y caer gravemente enferma. (**CONTRAER**)
473. Se administraron 600 mg de hidroxiclороquina (marca Plaquenil) a estos pacientes todos los días \*durante 10 días\*. (**ADMINISTRAR**)
474. Por otra parte, para octubre de 2020 y 9370 millones de personas vacunadas en segmento del 59% de la población mundial, \*la Organización Mundial de la Salud\* estimaba que al menos el 10 % de la población mundial ya se había contagiado de esta enfermedad (unas 780 millones de personas infectadas aproximadamente), debido al gran subregistro de casos a nivel mundial. (**ESTIMAR**)
475. \*Los médicos\* pueden determinar qué medidas tomar para cuidar al paciente, y pueden evitar que la condición del paciente empeore, resalta uno de los líderes de la investigación, Peter Kuhn. (**DETERMINAR**)
476. Todas las personas son susceptibles de adquirir el virus, siendo las que tienen más riesgo aquellas que tratan \*con personas contagiadas\*. (**TRATAR**)
477. Esta afección, además, es una señal de que tus pulmones podrían estar sufriendo severamente \*la infección con el SARS-CoV-2\*. (**SUFRIR**)
478. Desde el 24 de marzo de 2020 esta prueba pasó a ser gratuita para quienes pertenecen \*a cualquier tramo. (A, B, C o D)\*. (**PERTENECER**)
479. Si no tienes seguro, puedes comenzar con el departamento de salud estatal o los centros de salud comunitarios locales, que se conocen oficialmente como centros de salud calificados federalmente, aconsejó \*Harris\*, y agregó que algunos estados tienen un número de teléfono directo 1-800 para llamar. (**ACONSEJAR**)
480. El menor de 15 años manifestó que sí se quiere \*vacunar\* alegando que si se suministra es porque es confiable. (**QUERER**)
481. \*Los científicos\* evalúan la seguridad de los regímenes potenciados mixtos y controlan los niveles de anticuerpos resultantes. (**CONTROLAR**)
482. El gobierno federal está proporcionando vacunas contra el COVID-19 autorizadas por la FDA de forma gratuita a \*todas las personas que\* viven en los Estados Unidos, independientemente de su estado migratorio o de seguro médico. (**VIVIR**)
483. \*Desde el día 2 de marzo\* se incluye información sobre la evolución de la vacunación; los datos de vacunación no se actualizan los fines de semana y festivos. (**INCLUIR**)
484. Se han enviado \*suministros de desinfección, respiradores purificadores de aire, protectores faciales, guantes, reactivos de detección de ácido nucleico de coronavirus, ventiladores, monitores de pacientes, bombas de jeringa, bombas de infusión y alimentos a las áreas afectadas\*, e incluso han ofrecido pruebas gratuitas de COVID-19 al público. (**ENVIAR**)
485. El 31 de diciembre de 2019, \*la Organización Mundial de la Salud (OMS)\* recibió reportes de presencia de neumonía, de origen desconocido, en la ciudad de Wuhan, en China. (**RECIBIR**)
486. Separa \*ropa de cama y utensilios para el uso exclusivo de la persona enferma\*. (**SEPARAR**)

487. \*El SARS-CoV-2\* puede sobrevivir y permanecer contagioso en superficies inanimadas como metal, vidrio o plástico por varios días. **(PERMANECER)**
488. Esto causa \*irritación\* y por ello empezamos a toser, señala Schaffner. **(CAUSAR)**
489. Las rutas de transmisión de persona a persona del agente etiológico SARS-CoV-2 incluyen la transmisión directa por inhalación de microgotas y aerosoles liberadas a través de tos, estornudos, la respiración, el habla, los gritos o el canto,<sup>2</sup> o por contacto de las manos con superficies contaminadas, que \*luego\* tocan las membranas mucosas orales, nasales u oculares. **(TOCAR)**
490. Sin embargo, los adultos mayores y las personas de cualquier edad que tengan afecciones médicas previas deben llamar a su médico \*no bien aparecen los síntomas\*. **(LLAMAR)**
491. La variante ómicron del COVID-19 se expande rápidamente por el mundo y el Perú no es la excepción, los estudios apuntan \*a que es la más contagiosa hasta la fecha\*. **(APUNTAR)**
492. Presentaron linfopenia \*26 de 41 pacientes (63 %)\*. **(PRESENTAR)**
493. En un periodo donde las medidas impuestas instan a quedarse en casa, la población dominicana no solo tiene que hacer frente a que no todas las viviendas están correctamente acondicionadas (aproximadamente la mitad de los hogares no cuentan con agua potable), sino que \*la violencia de género, que en años anteriores presentaba unos datos realmente alarmantes\*, podría incrementarse. **(INCREMENTAR)**
494. Existe evidencia limitada que sugiere que el virus podría transmitirse uno o dos días antes de que se tengan síntomas, ya que la viremia alcanza un pico \*al final del período de incubación\*. **(ALCANZAR)**
495. Mantén las manos limpias y no te toques \*la cara\*. **(TOCAR)**
496. En este punto es cuando empezamos a sentirnos mal y perdemos \*el apetito\*. **(PERDER)**
497. De acuerdo con estudios exhaustivos al respecto, sabemos que \*el SRAS-CoV\* se transmitió de la civeta al ser humano y que se ha producido transmisión del MERS-CoV del dromedario al ser humano. **(TRANSMITIR)**
498. El proyecto Zoe también demostró \*que las personas que ya han sido vacunadas y dieron positivo estornudaban con más frecuencia\*. **(DEMOSTRAR)**
499. La prueba de antígeno también llamada prueba rápida da resultados \*en tan solo 15 minutos\*. **(DAR)**
500. Busca \*atención médica temprana\* y sigue las indicaciones de tu médica o médico, es importante conocer los síntomas de COVID-19 y los datos de alarma. **(BUSCAR)**
501. A comienzos de 2020, el mundo dirigió \*la mirada\* hacia China, donde el SARS-CoV-2 avanzaba por todo el país. **(DIRIGIR)**
502. Si la DSMB encuentra que el tratamiento experimental no está funcionando o está perjudicando a los participantes, la Junta recomendará que el estudio se detenga \*de inmediato\*. **(RECOMENDAR)**
503. Los estudios y los datos actuales muestran que \*los anticuerpos que nuestros cuerpos producen después de la vacunación\* reconocen y protegen contra las variantes del COVID-19. **(PROTEGER)**
504. \*El virus\* puede pasar a través de las mucosas, especialmente la mucosa nasal y laríngea, luego ingresa a los pulmones a través del tracto respiratorio. **(INGRESAR)**
505. \*Las mascarillas normales\* buscan principalmente reducir la propagación de la enfermedad evitando contagiar a otros, así que pueden dar una falsa sensación de seguridad a quienes las lleven. **(DAR)**
506. Cuando los cines y las salas de cine cerraron, \*la taquilla global\* cayó en miles de millones de dólares, mientras que el streaming se hizo más popular. **(CAER)**
507. \*En China\* se observó también que la mortalidad aumentaba con la edad y es más alta entre los hombres que en las mujeres (4,7 % frente al 2,8 %). **(AUMENTAR)**
508. \*Su proveedor\* puede tomar muestras con hisopos de la parte posterior de su nariz y garganta. **(TOMAR)**

509. Debe permanecer \*en casa\* y evitar el contacto con personas hasta que su proveedor le indique que es seguro terminar el aislamiento en casa. **(TERMINAR)**
510. En febrero de 2020, \*Aylward\* dirigió una delegación de la OMS a Wuhan por la crisis del coronavirus. **(DIRIGIR)**
511. \*Los fabricantes de vacunas\* envían actualizaciones de seguridad mensuales a la FDA. **(ENVIAR)**
512. Sanidad aprueba una cuarta dosis de la vacuna para personas inmunodeprimidas \*cinco meses después de la tercera\* **(APROBAR)**
513. España es el segundo por número de contagios, aunque \*ya\* empieza a ver la luz al final del túnel. **(VER)**
514. Sin embargo, ahora se están realizando \*estudios clínicos\* para probar la seguridad y eficacia de las vacunas disponibles actualmente en niños más pequeños. **(REALIZAR)**
515. \*Hacerse pruebas\* también proporciona información importante sobre el movimiento del virus dentro y entre comunidades. **(PROPORCIONAR)**
516. \*En este estudio\* se están evaluando las diferentes dosis de esta vacuna experimental para la seguridad y su capacidad de inducir una respuesta inmune en los voluntarios. **(EVALUAR)**
517. ¿O te estás cuidando \*a ti mismo\* en casa? **(CUIDAR)**
518. Los estudios están analizando si los tratamientos acortan el tiempo de recuperación y disminuyen \*los problemas respiratorios y de otro tipo de la infección por COVID-19\*. **(DISMINUIR)**
519. Para febrero de 2021 diez vacunas han sido autorizadas para su uso público \*por al menos una autoridad reguladora competente\*. **(AUTORIZAR)**
520. Según la Superintendencia de Salud, las Isapres al ser parte de la seguridad social y de la protección del derecho a la salud, no podrán, \*en ningún caso\*, excluir la cobertura a la enfermedad aunque haya sido declarada pandemia. **(EXCLUIR)**
521. Así pues, las mujeres se enfrentan a una carga triple de trabajo que puede agravar \*las desigualdades preexistentes\*. **(AGRAVAR)**
522. Investigadores especializados en enfermedades infecciosas de la Universidad de Texas en Austin (Estados Unidos) y otras instituciones de Hong Kong, China y Francia han llegado a la conclusión de que es muy probable que el coronavirus SARS-CoV-2 se propagara más allá de Wuhan y otras ciudades antes de que \*las autoridades chinas\* pudieran activar las medidas de cuarentena. **(ACTIVAR)**
523. Estas pequeñas estructuras son las que normalmente se llenan de aire, y a través de sus paredes se produce el intercambio gaseoso por el cual el oxígeno llega \*a la sangre\*, y de allí al resto del cuerpo. **(LLEGAR)**
524. Las mascarillas normales buscan principalmente reducir la propagación de la enfermedad evitando contagiar a otros, así que pueden dar una falsa sensación de seguridad \*a quienes las lleven\*. **(DAR)**
525. Asimismo, las medidas de aislamiento se mantendrán hasta que se hayan cumplido 10 días desde el inicio de los síntomas y, siempre y cuando \*estos síntomas\* hayan desaparecido. **(DESAPARECER)**
526. Para prevenir la expansión del virus, los gobiernos han impuesto \*restricciones de viajes, cuarentenas, confinamientos, aislamiento social, cancelación de eventos, y cierre de establecimientos\*. **(IMPONER)**
527. Unos investigadores han creado un algoritmo que permite \*a los relojes inteligentes\* tener la capacidad de detectar si tiene síntomas de la Covid-19. **(PERMITIR)**
528. El científico Vaclav Smil escribió \*en su libro Global Catastrophes and Trends (2008)\*, que la posibilidad de una pandemia de influenza durante los próximos 50 años es virtualmente del 100 %. **(ESCRIBIR)**
529. Según estudios que se han hecho sobre el SARS (síndrome respiratorio agudo grave), primo del covid-19, pensamos que puede ir \*a otras partes del cuerpo\*, dice Schaffner. **(IR)**
530. Al estornudar, cúbre\*te\* con la parte interna del codo. **(CUBRIR)**

531. Si estás cuidando a alguien con COVID-19 y no estás completamente vacunado, \*los CDC\* recomiendan que hagas cuarentena durante 14 días después del último contacto con la persona enferma y que estés atento a cualquier síntoma de la COVID-19. **(RECOMENDAR)**
532. Este gráfico de barras nos muestra \*el conteo acumulado de defunciones por COVID-19\*. **(MOSTRAR)**
533. Con esta información se puede obtener \*gráficos\* para comparar de manera visual el modo en que se ha visto impactada la movilidad a raíz de la pandemia. **(OBTENER)**
534. \*Los síntomas\* pueden durar unos días, y quienes tienen el virus podrían sentirse mejor en más o menos una semana. **(DURAR)**
535. \*El cuidador\* debe usar una mascarilla facial cuando esté cerca de la persona enferma, así como guantes si entra en contacto con su sangre, heces o fluidos corporales. **(USAR)**
536. \*A comienzos de 2020\*, el mundo dirigió la mirada hacia China, donde el SARS-CoV-2 avanzaba por todo el país. **(DIRIGIR)**
537. La Casa Blanca anunció que \*todos los viajeros internacionales que\* ingresan a los Estados Unidos deberán estar vacunados y este requisito comenzará a regir a partir del 8 de noviembre de 2021. **(INGRESAR)**
538. \*En diciembre de 2019\* hubo un brote epidémico de neumonía de causa desconocida en Wuhan, provincia de Hubei, China; el cual, según afirmó más tarde Reporteros Sin Fronteras, llegó a afectar a más de 60 personas el día 20 de ese mes. **(HABER)**
539. \*La revista The Lancet\* publicó el 24 de enero de 2020 un estudio de los primeros 41 casos de pacientes ingresados (en hospital de Wuhan) con el diagnóstico confirmado, desde el 16 de diciembre de 2019 al 2 de enero de 2020. **(PUBLICAR)**
540. Existe \*evidencia limitada que\* sugiere que el virus podría transmitirse uno o dos días antes de que se tengan síntomas, ya que la viremia alcanza un pico al final del período de incubación. **(SUGERIR)**
541. Sin embargo, \*usted\* puede propagar el virus incluso cuando no tiene síntomas. **(PROPAGAR)**
542. La Organización Mundial de la Salud (OMS) y la Organización Panamericana de la Salud OPS han reconocido el estado de pandemia, que supone \*el desarrollo de esta enfermedad\*. **(SUPONER)**
543. \*Algunos pacientes\* permanecen en el hospital por más tiempo. **(PERMANECER)**
544. Son varios los síntomas de la covid-19, pero hasta ahora no se ha conocido \*la ruta más probable de sintomatología\*. **(CONOCER)**
545. Su proveedor puede tomar muestras con hisopos \*de la parte posterior de su nariz y garganta\*. **(TOMAR)**
546. Hemos cerrado \*la sesión más antigua\* para que sigas navegando sin límites en el resto. **(CERRAR)**
547. Sin embargo, usted puede propagar el virus incluso \*cuando no tiene síntomas\*. **(PROPAGAR)**
548. \*Más de 19.000 profesores y 100.000 alumnos\* permanecieron esta semana en cuarentena por covid. **(PERMANECER)**
549. Los NIH establecieron \*la Red de Prevención de COVID-19 (COVPN)\* para coordinar las redes de investigación existentes, a fin de llevar a cabo grandes estudios clínicos en decenas de miles de personas de manera eficiente. **(ESTABLECER)**
550. Comparte esta información \*con tu familia, tus amigos y tus colegas\* para conseguir que todo el mundo conozca todos los datos relativos al COVID-19 y puedan proteger su salud. **(COMPARTIR)**
551. \*Otros tratamientos que se consideraron prometedores durante los primeros meses de la pandemia, como la hidroxiclороquina y el lopinavir/ritonavir\*, resultaron después ser inefectivos o incluso perjudiciales. **(RESULTAR)**
552. Otros han registrado \*la pérdida súbita del olfato y el gusto\* de forma menos frecuente (sin que la mucosidad fuese la causa). **(REGISTRAR)**
553. La gripe pertenece \*a la familia de la ortomixovirus\*, en que la que también se ubican la influenza estacionaria que se presenta cada año y puede cobrar vidas por sí misma. **(PERTENECER)**

554. El estudio, que observó \*los casos de más de 181 pacientes\*, concluyó que el tiempo medio de incubación es de 5 días, y que los síntomas pueden, por lo menores a largarse, otros 12 días. **(OBSERVAR)**
555. No comparta \*artículos personales como tazas, utensilios para comer, toallas o ropa de cama\*. **(COMPARTIR)**
556. Los expertos de la OMS explican \*la ciencia relacionada con la COVID-19\* en cada episodio. **(EXPLICAR)**
557. El virus puede pasar a través de las mucosas, especialmente la mucosa nasal y laríngea, luego ingresa \*a los pulmones\* a través del tracto respiratorio. **(INGRESAR)**
558. El Centro Nacional de Desarrollo Biotecnológico de China afirmó el 17 de marzo \*que el antiviral Favipiravir, un inhibidor de la ARN polimerasa, mostró resultados positivos en un estudio de casos y controles con 80 pacientes en el Hospital Popular n.o 3 de Shenzhen, los que recibieron tratamiento con Favipiravir dieron negativo dentro de un periodo más reducido de tiempo en comparación con los del grupo de control, y recomienda que se incluya en el tratamiento\*. **(AFIRMAR)**
559. Para reducir las posibilidades de infectarse, las organizaciones sanitarias recomiendan evitar el contacto cercano con personas enfermas; lavarse \*las manos\* frecuentemente con agua y jabón; no tocarse los ojos, la nariz o la boca con las manos sin lavar; y practicar una buena higiene respiratoria. **(LAVAR)**
560. Por el momento, no se recomienda ningún medicamento específico para prevenir o tratar la infección, pero los médicos de medio mundo trabajan a contrarreloj \*en desarrollar una cura\* tomando como referencia las experiencias pasadas. **(TRABAJAR)**
561. Los datos pasan por un proceso que requiere que el personal del Departamento de Salud de Puerto Rico revise, importe y asigne \*casos\*. **(REVISAR)**
562. Evite \*el transporte público, los vehículos compartidos y los taxis\*. **(EVITAR)**
563. Sin embargo, los adultos mayores y las personas de cualquier edad que tengan afecciones médicas previas deben llamar \*a su médico\* no bien aparecen los síntomas. **(LLAMAR)**
564. Pudimos determinar que los pacientes que no habían recibido Plaquenil (\*el medicamento que\* contiene hidroxyclorequina) seguían siendo contagiosos después de seis días, pero de los que habían recibido Plaquenil, después de seis días, solo el 25% seguía siendo contagioso. **(CONTENER)**
565. Así, los firmantes piden \*que se amplíe el cuadro clínico\* porque si no, el país podría pasar por alto o retrasar la identificación de muchos casos de Covid. **(PEDIR)**
566. La mayoría de las personas con la enfermedad tienen \*síntomas leves a moderados\* y se recuperan por completo. **(TENER)**
567. Las complicaciones incluían \*el síndrome de dificultad respiratoria aguda (resultado positivo para RT-PCR en tiempo real en la muestra de plasma), lesión cardíaca aguda, e infección secundaria\*. **(INCLUIR)**
568. Esto puede afectar tu capacidad para comer, dormir y concentrarte, así como empeorar \*problemas crónicos de salud\*. **(EMPEORAR)**
569. La persona que atienda \*al enfermo\* debe usar mascarilla y lavarse las manos tras cada contacto. **(ATENDER)**
570. \*Esto\* significa que la enfermedad se está extendiendo en varios países de todo el mundo, al mismo tiempo afectando a cientos de miles de personas en todos los continentes, causando un número de muertes que, a la fecha, sigue en ascenso. **(SIGNIFICAR)**
571. El 23 de agosto de 2021, la FDA aprobó la vacuna Pfizer-BioNTech, llamada Comirnaty, \*para personas mayores de 16 años\*. **(APROBAR)**
572. El virólogo y experto en influenza Robert Webster escribió en su libro Flu Hunter Unlocking the secrets of a virus (2019) \*que una pandemia mortal y disruptiva era solo cuestión de tiempo, y que la naturaleza eventualmente va a desafiar de nuevo a la humanidad con un equivalente al virus de la influenza de 1918\*. **(ESCRIBIR)**

573. En diciembre de 2019 hubo un brote epidémico de neumonía de causa desconocida \*en Wuhan, provincia de Hubei, China\*; el cual, según afirmó más tarde Reporteros Sin Fronteras, llegó a afectar a más de 60 personas el día 20 de ese mes. **(HABER)**
574. El 18 de marzo \*un artículo\* informa que el tratamiento con lopinavir/ritonavir da negativo en pruebas clínicas con 199 pacientes en China. **(INFORMAR)**
575. El espectro clínico de COVID-19 varía de formas asintomáticas a condiciones clínicas caracterizadas por \*insuficiencia respiratoria severa que\* requiere ventilación mecánica y apoyo en una unidad de cuidados intensivos (UCI), a manifestaciones en varios órganos y sistemas produciendo sepsis, shock séptico y síndrome de disfunción multiorgánica. **(REQUERIR)**
576. Un baño tibio o un baño de esponja puede ayudar \*a bajar la fiebre\*. **(AYUDAR)**
577. En febrero de 2020, Aylward dirigió \*una delegación de la OMS\* a Wuhan por la crisis del coronavirus. **(DIRIGIR)**
578. Sorprendentemente, \*las complicaciones trombóticas\* apenas se han descrito. **(DESCRIBIR)**
579. Se presentan los casos acumulados que se han reportado en Puerto Rico \*desde el comienzo de la epidemia del SARS-CoV-2\*. **(REPORTAR)**
580. Y esto ha llevado a que muchos gobiernos impongan \*nuevas restricciones\*. **(IMPONER)**
581. Los científicos pudieron advertir también \*que la anosmia -o pérdida del olfato- ya no se encuentra entre los 10 síntomas más comunes\*. **(ADVERTIR)**
582. Pon los guantes y las máscaras usados en un bote de basura con tapa \*en el cuarto de la persona enferma\*. **(PONER)**
583. De \*84 personas a las que\* se les diagnosticó Covid durante el estudio, el algoritmo detectó 67. **(DIAGNOSTICAR)**
584. Tong repitió la pregunta, momento en el que Aylward terminó \*la llamada\*. **(TERMINAR)**
585. \*Por el contrario, pueden agravar la condición del paciente\*, aconsejó. **(ACONSEJAR)**
586. Un par de meses después, Europa se convertía \*en el foco de la pandemia\*, y en mayo, América era el continente más afectado del mundo. **(CONVERTIR)**
587. \*Unos investigadores\* han creado un algoritmo que permite a los relojes inteligentes tener la capacidad de detectar si tiene síntomas de la Covid-19. **(CREAR)**
588. España es el segundo por número de contagios, aunque ya empieza a ver la luz \*al final del túnel\*. **(VER)**
589. La prueba de antígeno también llamada prueba rápida da \*resultados\* en tan solo 15 minutos. **(DAR)**
590. A finales de marzo, \*Estados Unidos\* se convirtió en el país más afectado del mundo. **(CONVERTIR)**
591. \*Las personas de cualquier edad que\* tengan fiebre o tos y además respiren con dificultad, sientan dolor u opresión en el pecho o tengan dificultades para hablar o moverse deben solicitar atención médica inmediatamente. **(RESPIRAR)**
592. La fatiga puede continuar \*mucho después de que el virus desaparezca\*. **(CONTINUAR)**
593. Después de que \*los investigadores\* prueban a fondo nuevos tratamientos o vacunas en el laboratorio para asegurarse de que puedan beneficiar a las personas, los tratamientos más prometedores pasan a los estudios clínicos. **(PROBAR)**
594. Tras analizar miles de casos de pacientes, durante el brote en China a principios de este año, los hospitales han determinado \*que existe un patrón de síntomas entre quienes se contagian de coronavirus\*. **(DETERMINAR)**
595. \*Este artículo\* trata sobre la descripción médica de la enfermedad. **(TRATAR)**
596. Una vez que las copias están listas, salen \*de la célula\* donde se originaron, la destruyen y comienzan a infectar a otras células. **(SALIR)**
597. En cualquier caso, aquellos que porten este material deben aprender a usarlo y eliminar \*lo\* correctamente. **(ELIMINAR)**
598. Casi la totalidad de instituciones educativas de todos los niveles cancelaron \*sus actividades presenciales\*, en cumplimiento de las disposiciones adoptadas por las autoridades de los distintos

- países, tendientes a disminuir la propagación de la pandemia y sus consecuencias fatales. **(CANCELAR)**
599. \*La RTHK\* respondió revisando el programa y no encontró ninguna violación de su carta. **(RESPONDER)**
600. COVID-19 es una enfermedad altamente infecciosa que suele producir \*síntomas que\* una persona puede controlar en casa. **(CONTROLAR)**
601. Si no es posible ponerte una mascarilla, cúbrete \*la boca y la nariz\* con un pañuelo o con el codo cuando tosas o estornudes. **(CUBRIR)**
602. Asegúrate \*de que en casa todos conocen cómo hacer un correcto lavado de manos\*. **(ASEGURAR)**
603. Aunque aún quedan muchas dudas sobre este nuevo virus que apareció por primera vez en la provincia central de Hubei en China hacia finales de 2019, se estima \*que el contagio se produce cuando aspiramos pequeñas gotas expulsadas a través de la tos o el estornudo de una persona infectada\*. **(ESTIMAR)**
604. \*A lo largo de todo el 2020 y en los meses transcurridos de 2021\* la información disponible acerca del coronavirus y su forma de transmisión fue variando. **(VARIAR)**
605. \*En Argentina, Canadá, Francia, Alemania, España, Reino Unido y Estados Unidos\*, las autoridades notificaron un incremento de las denuncias de violencia en el hogar. **(NOTIFICAR)**
606. Las mujeres menores de 50 años deben ser conscientes del raro riesgo de que se formen \*coágulos de sangre\* después de la vacunación\*. **(FORMAR)**
607. Sus síntomas (dolor de garganta y fiebre), pueden confundirse con los del coronavirus y hacer que nos alarmemos y acudamos \*al médico de cabecera\*. **(ACUDIR)**
608. Como el nombre oficial de la enfermedad se estableció después de la creación de este material, cualquier mención de nCoV se refiere \*a la COVID-19\*, la enfermedad infecciosa causada por el coronavirus descubierto más recientemente. **(REFERIR)**
609. Conforme a los hallazgos de la investigación británica, 1 de cada 5 pacientes experimenta también \*dolores\* en la zona abdominal. **(EXPERIMENTAR)**
610. Si no dispones de agua ni jabón, usa \*un desinfectante de manos que\* contenga por lo menos un 60% de alcohol. **(CONTENER)**
611. Al toser o estornudar, cubrirse la boca y la nariz con la sangría o fosa cubital (la concavidad que forma \*la cara interna del brazo\* al flexionarlo por el codo). **(FORMAR)**
612. Hay \*casos confirmados de coronavirus\* en casi todo del mundo, incluyendo China, Europa (siendo en Italia y España los primeros focos), Irán, Estados Unidos, África, Australia y LATAM. **(HABER)**
613. Que los animales de compañía propaguen el nuevo coronavirus ha sido también negado por la institución, que no tiene ninguna prueba de que el nuevo coronavirus infecte \*a animales de compañía como los perros y los gatos\*. **(INFECTAR)**
614. Tal y como han informado los Institutos Nacionales de Salud (NIH, en inglés) a través de un comunicado, \*la fase 1 de la prueba clínica para evaluar una vacuna en investigación diseñada frente a la enfermedad del coronavirus\* ha empezado en el Instituto de Investigación Sanitaria Kaiser Permanente Washington, en Seattle, y se prevé que se alargue hasta seis semanas. **(EMPEZAR)**
615. \*Alrededor del 39 %\* se enfermó moderadamente, con síntomas adicionales que incluyen neumonía o problemas pulmonares revelados por la tomografía computarizada, pero sin dificultad respiratoria evidente. **(ENFERMAR)**
616. Los médicos pueden enviar \*a los pacientes leves\* a recuperarse en su casa, bajo las siguientes condiciones: **(ENVIAR)**
617. Y aunque esas consecuencias no resulten muy dañinas, lo cierto es que pueden ser muy duraderas, ya que los investigadores hablan, incluso, de la existencia de secuelas incluso diez semanas después de que \*el virus SARS-CoV-2\* haya desaparecido del organismo del paciente. **(DESAPARECER)**

618. El coronavirus afecta de distintas maneras a cada persona, pero existe un amplio abanico de afecciones a causa de la enfermedad \*en personas vacunadas\*(EXISTIR)
619. \*Los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés)\* ofrecen más información sobre la autoevaluación, por ejemplo, cómo hacerse un hisopado. (OFRECER)
620. Las citoquinas son proteínas solubles que actúan \*sobre las células del sistema inmune\* y que regulan la activación, proliferación y reclutamiento celular. (ACTUAR)
621. El virus puede pasar a través de las mucosas, especialmente la mucosa nasal y laríngea, luego ingresa a los pulmones \*a través del tracto respiratorio\*. (PASAR)
622. Se permiten visitas y acompañamiento \*a pacientes NO COVID\* en centros hospitalarios, bajo determinadas condiciones(PERMITIR)
623. Para reducir las posibilidades de infectarse, las organizaciones sanitarias recomiendan evitar el contacto cercano con personas enfermas; lavarse las manos \*frecuentemente\* con agua y jabón; no tocarse los ojos, la nariz o la boca con las manos sin lavar; y practicar una buena higiene respiratoria. (LAVAR)
624. El gobierno de Hong Kong advirtió \*a cualquiera que viaje fuera de la ciudad\* que no toque animales; no coma carne de caza; y evite visitar mercados húmedos, mercados de aves vivas y granjas. (ADVERTIR)
625. Cada vez es más complicado distinguir la covid del resto de infecciones por virus y a medida que van saliendo nuevas variantes parece que \*diferenciar una gripe o una bronquitis de la Covid\* no resulta tan sencillo. (RESULTAR)
626. La pandemia de la COVID-19 está en su segunda ola y ya ha dejado \*en nuestro país\* más de 900.000 personas infectadas. (DEJAR)
627. \*Muchas personas\* no presentan efectos secundarios. (PRESENTAR)
628. \*Las gotas en el aire\* también pueden caer sobre superficies y objetos. (CAER)
629. Es fundamental que la población no se relaje, no baje la guardia para que no perdamos lo que hemos logrado \*en los últimos meses\*: la reducción y estabilidad de la epidemia y su control asistencial. (LOGRAR)
630. \*Los estudios\* están analizando si los tratamientos acortan el tiempo de recuperación y disminuyen los problemas respiratorios y de otro tipo de la infección por COVID-19. (ANALIZAR)
631. \*El Ministerio de Defensa de China\* asegura haber desarrollado con éxito una vacuna contra el coronavirus y ha autorizado las pruebas en humanos. (AUTORIZAR)
632. Con el cierre de la actividad regular algunas productoras han cedido material a los hospitales y gran número de creadores, actores y bailarines han acudido \*al mundo digital, en especial a las redes sociales\*, para respaldar las diversas iniciativas sociales, reivindicar el sector o simplemente aportar nuevo contenido cultural y seguir haciendo arte. (ACUDIR)
633. La investigación clínica encontró que se detecta una alta concentración de citoquinas \*en el plasma de pacientes críticos infectados con SARS-CoV-2\*, lo que sugiere que la tormenta de citocinas se asociaba con la gravedad de la enfermedad. (ENCONTRAR)
634. La partícula viral viajará, a través del sistema de transporte de vesículas celular en el que interviene el aparato de Golgi, \*hasta la superficie\*. (VIAJAR)
635. En total, se estima que haber contraído el virus proporciona \*una inmunización natural contra los síntomas de COVID-19 del 94%\*. (PROPORCIONAR)
636. \*La COVID-19\* se está propagando rápidamente por todo el mundo y dentro de los Estados Unidos. (PROPAGAR)
637. Se procuró que la denominación no contuviera \*nombres de personas o referencias a ningún lugar, especie animal, tipo de comida, industria, cultura o grupo de personas\*, en línea con las recomendaciones internacionales, para evitar que hubiera estigmatización contra algún colectivo. (CONTENER)
638. Los científicos evalúan la seguridad de los regímenes potenciados mixtos y controlan \*los niveles de anticuerpos resultantes\*. (CONTROLAR)

639. Los adultos mayores y aquellos con afecciones crónicas graves se enfrentan a un mayor riesgo de enfermedades graves y complicaciones por COVID-19 y \*los CDC de Estados Unidos\* les han aconsejado que eviten las multitudes y se queden en casa tanto como sea posible en áreas de brote comunitario. **(ACONSEJAR)**
640. La FDA analizó los datos de los estudios clínicos \*de inmediato\*. **(ANALIZAR)**
641. \*En China\* se observó también que la mortalidad aumentaba con la edad y es más alta entre los hombres que en las mujeres (4,7 % frente al 2,8 %). **(OBSERVAR)**
642. \*El 14 de febrero\* murió un turista chino en París, en la que se creyó se trataba de la primera muerte por COVID-19 fuera de Asia pero posteriormente se descubrió que un hombre fallecido en la ciudad española de Valencia el 13 de febrero también estaba infectado de coronavirus, que había contraído en Nepal. **(MORIR)**
643. En general, \*el coronavirus\* afecta menos a la parte alta del aparato respiratorio, a diferencia del resfriado común aunque, como decimos, las nuevas variantes han cambiado este patrón. **(AFECTAR)**
644. \*Un informe de la Organización Mundial de la Salud\* incluye el dolor de garganta entre los síntomas que pueden padecer los pacientes con coronavirus. **(INCLUIR)**
645. La anosmia, en particular, se ha visto \*en pacientes que finalmente dieron positivo para el coronavirus sin otros síntomas\*, según la Academia Americana de Otorrinolaringología-Cirugía de Cabeza y Cuello. **(VER)**
646. Necesitamos solidarizarnos con las historias entrañables que nos cuentan \*en las redes\*, pero en la inmensidad de internet también necesitamos un recordatorio de que esa historia nos importó un día. **(CONTAR)**
647. \*Actualmente\*, estamos viviendo una tercera ola con datos preocupantes en todo el mundo y un diagnóstico prematuro puede evitar contagios de las personas más vulnerables al virus (personas mayores, afectados por patologías previas...) si se guarda de manera estricta el aislamiento de dos semanas posterior a la detección. **(VIVIR)**
648. Comparte \*tus preocupaciones\*. **(COMPARTIR)**
649. \*La OMS\* anunció el 11 de febrero de 2020 que COVID-19 sería el nombre oficial de la enfermedad. **(ANUNCIAR)**
650. El aislamiento ha provocado que muchos enfermos no puedan continuar \*con sus tratamientos\*, agravando sus problemas de salud. **(CONTINUAR)**
651. El espectro clínico de COVID-19 varía de formas asintomáticas a condiciones clínicas caracterizadas por insuficiencia respiratoria severa que requiere \*ventilación mecánica y apoyo\* en una unidad de cuidados intensivos (UCI), a manifestaciones en varios órganos y sistemas produciendo sepsis, shock séptico y síndrome de disfunción multiorgánica. **(REQUERIR)**
652. La OMS anunció el \*11 de febrero de 2020\* que COVID-19 sería el nombre oficial de la enfermedad. **(ANUNCIAR)**
653. La India registra \*casi 200.000 nuevos casos de coronavirus\* en las últimas 24 horas **(REGISTRAR)**
654. La radiología torácica (bien por radiografía, tomografía computerizada o ecografía) puede ayudar al diagnóstico de COVID-19 e identificar o descartar \*complicaciones pulmonares\*. **(DESCARTAR)**
655. El coronavirus se extendió \*por España\* desde febrero de 2020 y a mediados de marzo obligó a imponer el estado de alarma para tratar de frenar los contagios y aliviar la saturación en los hospitales. **(EXTENDER)**
656. La Organización Mundial de la Salud recomienda por este motivo que se desinfecten las superficies, especialmente \*en el entorno sanitario\*. **(RECOMENDAR)**
657. \*Actualmente (agosto 2021)\*, existe la recomendación de vacunar a toda mujer embarazada. **(EXISTIR)**
658. La gente no se está dando cuenta de este cambio y podría pensar \*que tiene algún tipo de resfriado estacional\*, van a fiestas y pueden contagiar. **(PENSAR)**

659. La información de la situación epidemiológica y la evolución de la infección por coronavirus se actualiza con datos oficiales aportados por los diferentes sistemas de salud autonómicos al Centro de Coordinación de Alertas y Emergencias Sanitarias (CCAES), donde se pueden consultar \*los datos de España, Europa y el resto del Mundo\*. **(CONSULTAR)**
660. En otros casos, \*el coronavirus\* provocó infartos y accidentes cerebrovasculares (ACV) en personas jóvenes e insensibilidad a los efectos de la falta de oxígeno, que se conoce como hipoxia feliz. **(PROVOCAR)**
661. \*La FDA\* analizó los datos de los estudios clínicos de inmediato. **(ANALIZAR)**
662. \*El virólogo y experto en influenza Robert Webster\* escribió en su libro Flu Hunter Unlocking the secrets of a virus (2019) que una pandemia mortal y disruptiva era solo cuestión de tiempo, y que la naturaleza eventualmente va a desafiar de nuevo a la humanidad con un equivalente al virus de la influenza de 1918. **(ESCRIBIR)**
663. \*En esta página\* se recoge documentación de la Comunidad de Madrid, el Estado, el Ayuntamiento de Madrid y la Unión Europea. **(RECOGER)**
664. \*España\* hasta el momento suma un total de 3.971.124 casos positivos de Covid-19 y ya son 81.020 las personas fallecidas, según los últimos datos proporcionados por el Ministerio de Sanidad. **(SUMAR)**
665. CCINSHAE y Sedesa llamaron a poner atención en seguir adelante con las medidas de seguridad que eviten \*la propagación de contagios\*. **(EVITAR)**
666. El 23 de agosto de 2021, \*la FDA\* aprobó la vacuna Pfizer-BioNTech, llamada Comirnaty, para personas mayores de 16 años. **(APROBAR)**
667. Los investigadores creen \*que la variante Delta proveniente de la India podría estar detrás del cambio en el cuadro clínico\***(CREER)**
668. \*Los estudios\* muestran que las vacunas contra el COVID-19 son muy efectivas para prevenir el COVID-19, incluso para personas con alto riesgo de contraer la enfermedad. **(MOSTRAR)**
669. El 10 de enero de 2020, las primeras secuencias del genoma completo del SARS-CoV-2 estuvieron disponibles en GISAID, \*lo que\* permitió respuestas globales a la pandemia, incluyendo el desarrollo de las primeras vacunas y pruebas de diagnóstico para detectar el SARS-CoV-2. **(PERMITIR)**
670. Los pacientes que pierden \*el olfato\* no siempre pierden el gusto. **(PERDER)**
671. \*El COVID-19\* ha demostrado, desde las primeras olas de la pandemia, que las enfermedades físicas también pueden causar afecciones mentales. **(DEMOSTRAR)**
672. ¿Por qué es importante quedarnos \*en casa\*? **(QUEDAR)**
673. Independientemente de \*las multas que\* pueden imponer las diferentes administraciones, los hechos podrían ser delito. **(IMPONER)**
674. La gripe suele ser la única que \*nos\* hace sentir dolor muscular. **(SENTIR)**
675. Debes tener en cuenta que esto ha cambiado mucho ahora, tras varios picos de infección que hemos pasado y \*las nuevas variantes que\* han surgido. **(SURGIR)**
676. La infección se produce cuando una persona enferma tose o estornuda y expulsa partículas del virus que entran en contacto \*con otras personas\*. **(ENTRAR)**
677. La pandemia de la COVID-19 está en su segunda ola y \*ya\* ha dejado en nuestro país más de 900.000 personas infectadas. **(DEJAR)**
678. Al principio, \*la ciencia\* no pensaba que la diarrea u otros problemas gástricos típicos que a menudo vienen con la gripe se aplicarían al nuevo coronavirus, también conocido como SARS-CoV-2. **(PENSAR)**
679. El brote de coronavirus ha sido ampliamente reportado \*en los noticieros del mundo\* y la Organización Mundial de la Salud (OMS), lo ha declarado oficialmente como una pandemia, desde el pasado 11 de marzo. **(REPORTAR)**
680. Rápidamente, a principios de enero, \*las autoridades de este país\* identificaron la causa como una nueva cepa de coronavirus. **(IDENTIFICAR)**

681. Los pacientes mayores y las personas que tienen afecciones médicas subyacentes o están inmunocomprometidos deben comunicarse con su médico \*temprano\* incluso en el curso de una enfermedad leve, aconseja el CDC. **(COMUNICAR)**
682. Las infecciones asintomáticas son las más comunes entre niños, pero las afecciones dependen de la edad, ya que el dolor estómago es un síntoma común en los más pequeños, mientras que los mayores de 12 años pueden sufrir una pérdida del gusto o del olfato, según reconocieron \*un grupo de expertos en epidemiología\*. **(RECONOCER)**
683. La RT-PCR da positivo durante varias semanas después de la primera infección (30 días de media, según algunos estudios), ya que detecta la presencia del ARN del virus, aunque este ya no sea viable y \*el paciente\* haya superado la infección y ya no sea contagioso. **(SUPERAR)**
684. Recientemente (diciembre 2021), se ha publicado una revisión sistemática en la que estudios observacionales en los que participaron casi dos millones de adultos sugieren que la deficiencia / insuficiencia de vitamina D aumenta \*la susceptibilidad al COVID-19 y al COVID-19 grave\*. **(AUMENTAR)**
685. \*Esta crisis sanitaria\* ha evidenciado algunos problemas en materia de igualdad de género que existen en este país, sobre todo en el ámbito laboral y en especial en el sector de cuidados. **(EVIDENCIAR)**
686. Si desarrolla cualquiera de los síntomas relacionado con el coronavirus, es importante que busque \*atención médica\* tan pronto como sea posible. **(BUSCAR)**
687. El gobierno de Hong Kong advirtió a cualquiera que viaje fuera de la ciudad \*que no toque animales; no coma carne de caza; y evite visitar mercados húmedos, mercados de aves vivas y granjas\*. **(ADVERTIR)**
688. Existe evidencia limitada que sugiere que el virus podría transmitirse uno o dos días antes de que se tengan síntomas, ya que la viremia alcanza \*un pico\* al final del período de incubación. **(ALCANZAR)**
689. \*La Organización\* apoya a los países en la vigilancia, la prevención y la preparación para el control de enfermedades pandémicas y propensas a epidemias a través del desarrollo de estrategias basadas en la evidencia para predecir, prevenir, detectar y responder. **(APOYAR)**
690. Desde principios de mayo, hemos estado observando \*los principales síntomas de los usuarios de la aplicación\*, y no son los mismos que antes, explica el experto. **(OBSERVAR)**
691. Estos estudios se centran en \*personas a las que\* se les diagnostica COVID-19 y están hospitalizadas, así como en aquellas que se están recuperando de COVID-19. **(DIAGNOSTICAR)**
692. Para reducir las posibilidades de infectarse, las organizaciones sanitarias recomiendan evitar el contacto cercano con personas enfermas; lavarse las manos frecuentemente \*con agua y jabón\*; no tocarse los ojos, la nariz o la boca con las manos sin lavar; y practicar una buena higiene respiratoria. **(LAVAR)**
693. A finales de mayo, América pasa a ser el epicentro mundial de la pandemia del coronavirus, pero la llegada de las vacunas en 2021 y su administración en los países más desarrollados ha hecho aparecer una mayor transmisión relativa de la COVID-19 \*en Asia\*. **(APARECER)**
694. Y, aunque \*calcular la letalidad con precisión\* puede resultar prematuro por la omisión de casos leves de los registros, el virus se ha cobrado vidas a un ritmo diferente en cada país. **(RESULTAR)**
695. Las personas deben seguir \*las directrices establecidas por sus organizaciones locales de salud\* para garantizar que se mantengan a salvo y mantengan así a los demás. **(SEGUIR)**
696. Los síntomas pueden durar unos días, y quienes tienen el virus podrían sentirse mejor \*en más o menos una semana\*. **(SENTIR)**
697. La enfermedad se propaga principalmente de persona a persona a través de las gotículas que salen despedidas de la nariz o la boca de una persona infectada \*al toser, estornudar o hablar\*. **(SALIR)**
698. \*Los CDC\* ofrecen una guía para el tratamiento de COVID-19 en el hogar. **(OFRECER)**
699. Quienes ya se estaban recuperando pueden ver que sus fiebres desaparecen \*después de 12 días\*. **(VER)**

700. Se administraron 600 mg de hidroxiclороquina (marca Plaquenil) a estos pacientes \*todos los días\* durante 10 días. **(ADMINISTRAR)**
701. Todas las vacunas pueden causar \*algunos efectos secundarios\*, que son signos normales de que su cuerpo está produciendo anticuerpos. **(CAUSAR)**
702. Jeremy Konyndyk, exdirector de la Oficina de USAID de Asistencia para Desastres Externos de Estados Unidos escribió en un artículo de 2017 \*que una nueva gran crisis sanitaria global es una cuestión de 'cuándo', no de 'sí'\*, y añadió que el virus sería comparable al de la pandemia de 1918. **(ESCRIBIR)**
703. La OMS actualiza \*los indicios\* para descubrir si estás contagiado; consulta en el gráfico todas las partes del cuerpo a las que puede afectar el virus **(ACTUALIZAR)**
704. \*Sus investigadores\* han analizado los síntomas vinculados a la variante delta y la nueva variante ómicron. **(ANALIZAR)**
705. A finales de enero, el Ministerio de Salud de Rusia identificó \*tres medicamentos para adultos que podrían ayudar a tratar la enfermedad\*. **(IDENTIFICAR)**
706. \*La pandemia de la COVID-19\* está en su segunda ola y ya ha dejado en nuestro país más de 900.000 personas infectadas. **(DEJAR)**
707. El virus puede pasar a través de las mucosas, especialmente la mucosa nasal y laríngea, \*luego\* ingresa a los pulmones a través del tracto respiratorio. **(INGRESAR)**
708. Schaffner compara \*la respuesta inflamatoria\* con un conflicto bélico. **(COMPARAR)**
709. Un estudio de más de 550 hospitales en China también identificó que \*los pacientes hospitalizados\* desarrollaron neumonía al tercer día de su enfermedad. **(DESARROLLAR)**
710. La variante ómicron se ha notificado \*hasta ahora\* en 19 países y territorios de la región. **(NOTIFICAR)**
711. \*Los expertos\* esperan que las vacunas estén autorizadas para niños y adolescentes a fines de 2021. **(ESPERAR)**
712. No reutilice máscaras de un solo uso; deseche las máscaras de un solo uso \*después de cada uso\* y deséchelas inmediatamente después de retirarlas. **(DESECHAR)**
713. Debe permanecer en casa y evitar el contacto con personas hasta que su proveedor le indique \*que es seguro terminar el aislamiento en casa\*. **(INDICAR)**
714. \*Todo el mundo\* habla de la enfermedad por coronavirus 2019 (COVID-19), y allá donde mires hay información sobre el virus y sobre cómo protegerte. **(HABLAR)**
715. Esto se debe sobre todo a la dependencia económica y a las dificultades para denunciar que han provocado e incrementado \*las nuevas medidas de confinamiento y aislamiento social\*. **(INCREMENTAR)**
716. \*Este gráfico de barras\* nos muestra el conteo acumulado de defunciones por COVID-19. **(MOSTRAR)**
717. En China se observó también que \*la mortalidad\* aumentaba con la edad y es más alta entre los hombres que en las mujeres (4,7 % frente al 2,8 %). **(AUMENTAR)**
718. La gente no se está dando cuenta de este cambio y podría pensar que tiene algún tipo de resfriado estacional, van \*a fiestas\* y pueden contagiar. **(IR)**
719. Los fabricantes de vacunas envían \*actualizaciones de seguridad mensuales\* a la FDA. **(ENVIAR)**
720. Los CDC han actualizado las recomendaciones para el público sobre el aislamiento y la cuarentena y están revisando \*el sitio web de los CDC\* para incluir estos cambios. **(REVISAR)**
721. \*Facebook\* comunicó que eliminaría todo contenido que hiciera apología a las protestas contra la inmovilización social, el propio Mark Zuckerberg calificó las difusiones de las protestas por redes sociales de información errónea dañina. **(ELIMINAR)**
722. Mantener al menos un metro de distancia de otras personas, particularmente \*aquellas que\* tosan, estornuden y tengan fiebre. **(TOSER)**
723. Según estudios que se han hecho sobre \*el SARS (síndrome respiratorio agudo grave), primo del covid-19\*, pensamos que puede ir a otras partes del cuerpo, dice Schaffner. **(IR)**
724. Muchas personas no presentan \*efectos secundarios\*. **(PRESENTAR)**

725. Los padres y cuidadores de niños que presentan señales de una infección por coronavirus deben comunicarse \*con su médico\* y mantener al niño en casa. **(COMUNICAR)**
726. Sin embargo, la mayoría de los niños y adultos no serán considerados febriles hasta que \*su temperatura\* alcance los 100 grados Fahrenheit (37,7 grados Celsius). **(ALCANZAR)**
727. Permanezca \*en su domicilio\* y contacte por teléfono con su centro de salud, a través del teléfono habitual de cita, donde le darán las indicaciones necesarias. **(PERMANECER)**
728. El coronavirus provoca \*pérdida de gusto y olfato, diarrea\*... **(PROVOCAR)**
729. Uno de los misterios que aún rodean a este coronavirus está en su capacidad de mutar, algo que será clave para entender si bastará con una vacuna para erradicarlo o si por el contrario volverá \*cada nueva temporada\* con una forma distinta. **(VOLVER)**
730. Así son los síntomas del coronavirus día a día, un desglose muestra \*la evolución de cada uno\*. **(MOSTRAR)**
731. Para que la alarma saltara, la frecuencia cardíaca debía elevarse \*durante un espacio prolongado de tiempo\*, por lo que si alguno salía a correr no saltaba. **(ELEVAR)**
732. Ya hemos hablado \*de China\*. **(HABLAR)**
733. \*Los científicos de los NIH\* han estado estudiando las vacunas de ARNm y los coronavirus durante décadas. **(ESTUDIAR)**
734. De 84 personas a las que se les diagnosticó \*Covid\* durante el estudio, el algoritmo detectó 67. **(DIAGNOSTICAR)**
735. Lavarse \*las manos\* con frecuencia durante 20 segundos con agua y jabón es fundamental para el cuidador. **(LAVAR)**
736. Las máscaras limitan \*el volumen y la distancia de viaje de las gotas respiratorias dispersas\* al hablar, estornudar y toser. **(LIMITAR)**
737. \*Las personas\* también pueden aplicar métodos de distanciamiento social limitando los viajes, evitando áreas llenas de gente y alejándose físicamente de las personas enfermas. **(APLICAR)**
738. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió \*expertos\* a Wuhan para apoyar la investigación. **(ENVIAR)**
739. \*Todas las vacunas\* pueden causar algunos efectos secundarios, que son signos normales de que su cuerpo está produciendo anticuerpos. **(CAUSAR)**
740. El equipo encontró más casos relacionados al mercado y el 30 de diciembre \*las autoridades de salud de Wuhan\* comunicaron los casos al CCDC, que envió expertos a Wuhan para apoyar la investigación. **(COMUNICAR)**
741. Según información de la OMS, \*el coronavirus\* se transmite por contacto de persona a persona con algún infectado. (incluso si no presenta síntomas). **(TRANSMITIR)**
742. ¿O \*te\* estás cuidando a ti mismo en casa? **(CUIDAR)**
743. \*El equipo\* encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió expertos a Wuhan para apoyar la investigación. **(ENCONTRAR)**
744. Esto significa \*cubrirte la boca y la nariz con el codo o pañuelo doblado cuando toses o estornudas\*. **(SIGNIFICAR)**
745. Las pruebas de laboratorio para el diagnóstico de la COVID-19 (enfermedad causada por el SARS-CoV-2) incluyen \*los métodos que\* detectan la presencia del propio virus y aquellos que detectan los anticuerpos producidos por el cuerpo humano en respuesta a la infección. **(DETECTAR)**
746. El coronavirus afecta de distintas maneras a cada persona, pero existe \*un amplio abanico de afecciones a causa de la enfermedad\* en personas vacunadas **(EXISTIR)**
747. Si se cuenta \*con una vacuna o esquema completo\* y se presentan síntomas como tos, ronquera, dolor de cabeza y/o garganta, escurrimiento nasal y cuerpo cortado, se debe asumir que es covid-19, aislarse de manera inmediata y llamar a Locatel (55-5658-1111). **(CONTAR)**
748. El dolor de cabeza de la COVID suele tender a empeorar si \*la persona\* hace esfuerzos o movimientos de cabeza. **(HACER)**

749. Dra. Nora Volkow habla \*sobre un estudio que muestra que las personas con trastornos por consumo de sustancias, incluido el trastorno por consumo de marihuana, pueden tener un mayor riesgo de contraer infecciones irruptivas del SARS-CoV-2\*. **(HABLAR)**
750. Los estudios realizados en Reino Unido, de seguimiento de síntomas desde los inicios de la pandemia, evidenciaron durante la ola de delta \*que los dolores de cabeza se habían convertido en el síntoma más común de la infección con la variante\*. **(EVIDENCIAR)**
751. (HHS) está desarrollando \*un plan\* para distribuir dosis adicionales de vacunas de ARNm a quienes califiquen para recibirlas. **(DESARROLLAR)**
752. Limpie todas las áreas de contacto frecuente \*en el hogar\*, como las manijas de las puertas, accesorios de baño y cocina, inodoros, teléfonos, tabletas, encimeras y otras superficies. **(LIMPIAR)**
753. Una mujer china de 61 años vecina de Wuhan que el 8 de enero viajó en vuelo directo \*de Wuhan\* a Tailandia en compañía de cinco miembros de su familia en un grupo de dieciséis personas. **(VIAJAR)**
754. \*Esta simplificación\* se explicaría por el fenómeno de la economía lingüística, especialmente en el registro informal y discurso hablado. **(EXPLICAR)**
755. El test debe realizarse \*antes de acudir a la reunión\*, dado que el resultado de la prueba solo tarda 15 minutos. **(REALIZAR)**
756. Tong repitió la pregunta, momento en el que \*Aylward\* terminó la llamada. **(TERMINAR)**
757. También se recupera \*la autorización de baño en ríos, lagos y pozas remansadas de agua dulce\*. **(RECUPERAR)**
758. También es frecuente el dolor de cabeza, la fiebre, el olor reducido, la respiración fuerte, el dolor abdominal y el dolor muscular, aunque, por lo general, suelen desaparecer \*pasados dos días\*. **(DESAPARECER)**
759. \*El brote de coronavirus\* ha sido ampliamente reportado en los noticieros del mundo y la Organización Mundial de la Salud (OMS), lo ha declarado oficialmente como una pandemia, desde el pasado 11 de marzo. **(REPORTAR)**
760. La investigación clínica encontró que se detecta \*una alta concentración de citoquinas en el plasma de pacientes críticos infectados con SARS-CoV-2\*, lo que sugiere que la tormenta de citocinas se asociaba con la gravedad de la enfermedad. **(DETECTAR)**
761. El 9 de febrero de 2021, una comisión internacional de la OMS enviada a China para investigar \*los orígenes de la pandemia\* in situ concluyó que el foco inicial de la pandemia fue el determinado desde un principio, el mercado mayorista de Wuhan. **(INVESTIGAR)**
762. El virus infecta \*las vías respiratorias\* y causa síntomas que van desde un cuadro leve (tos seca, fiebre...) a insuficiencia respiratoria aguda y neumonías potencialmente mortales. **(INFECTAR)**
763. Los adultos mayores y aquellos con afecciones crónicas graves se enfrentan a un mayor riesgo de enfermedades graves y complicaciones por COVID-19 y los CDC de Estados Unidos \*les\* han aconsejado que eviten las multitudes y se queden en casa tanto como sea posible en áreas de brote comunitario. **(ACONSEJAR)**
764. \*La gripe\* suele ser la única que nos hace sentir dolor muscular. **(SENTIR)**
765. Tras una última fase de maduración, en la que intervienen proteasas virales, todos los componentes del virus encajarán, \*la partícula\* será infecciosa y podrá comenzar un nuevo ciclo celular. **(COMENZAR)**
766. La India registra casi 200.000 nuevos casos de coronavirus \*en las últimas 24 horas\*. **(REGISTRAR)**
767. En abril, se anunció \*que la ivermectina inhibe la replicación del SARS-CoV-2 in vitro\*. **(ANUNCIAR)**
768. \*Exitosos libros como A Dancing Matrix (1994) de Robin Marantz Henig, The Coming Plague (1994) de Laurie Garrett y The Hot Zone de Richard Preston (1995)\* tocaron este tema y coincidieron en que mundialmente no existía la preparación suficiente para lidiar con algo así. **(TOCAR)**

769. La Organización Mundial de la Salud (OMS) y la Organización Panamericana de la Salud OPS han reconocido \*el estado de pandemia, que\* supone el desarrollo de esta enfermedad. **(SUPONER)**
770. España \*hasta el momento\* suma un total de 3.971.124 casos positivos de Covid-19 y ya son 81.020 las personas fallecidas, según los últimos datos proporcionados por el Ministerio de Sanidad. **(SUMAR)**
771. Los NIH están patrocinando estudios para probar la seguridad y eficacia de los inmunomoduladores, que minimizan \*los efectos de una respuesta hiperactiva del sistema inmunológico que\* experimentan algunos pacientes con COVID-19. **(EXPERIMENTAR)**
772. Para febrero de 2021 \*diez vacunas\* han sido autorizadas para su uso público por al menos una autoridad reguladora competente. **(AUTORIZAR)**
773. \*Los CDC\* han actualizado las recomendaciones para el público sobre el aislamiento y la cuarentena y están revisando el sitio web de los CDC para incluir estos cambios. **(ACTUALIZAR)**
774. Es posible que también aumente \*tu consumo de alcohol, tabaco u otras sustancias\*. **(AUMENTAR)**
775. Jeremy Konyndyk, exdirector de la Oficina de USAID de Asistencia para Desastres Externos de Estados Unidos escribió \*en un artículo de 2017\* que una nueva gran crisis sanitaria global es una cuestión de 'cuándo', no de 'sí', y añadió que el virus sería comparable al de la pandemia de 1918. **(ESCRIBIR)**
776. Atendiendo a los datos del IMSERSO, \*las mujeres\* representan el 70,4% de las personas que viven en residencias, por un 29,6% de hombres. **(REPRESENTAR)**
777. Las máscaras quirúrgicas (normales) son utilizadas \*desde hace tiempo\* por personas sanas en Hong Kong, Japón, Singapur y Malasia. **(UTILIZAR)**
778. No obstante, actualmente parece ser que es más probable que una persona contagie cuando tiene síntomas, tales como fiebre o tos, pero se desconoce \*el periodo de tiempo durante el cual una persona con SARS puede transmitir la enfermedad a otras\*, es decir, podrían aparecer en tan solo 2 días. **(DESCONOCER)**
779. Las declaraciones emitidas por Xi Jinping el 3 de febrero, indicó la necesidad de un énfasis por parte de los medios estatales en contar las historias conmovedoras de cómo las personas en la línea del frente están previniendo y luchando contra \*el virus\* como una prioridad de cobertura, mientras que el alto funcionario Zhang Xiaoguo dijo que su departamento trataría la propaganda con respecto a las medidas de control y prevención del virus como su máxima prioridad. **(PREVENIR)**
780. \*El test\* debe realizarse antes de acudir a la reunión, dado que el resultado de la prueba solo tarda 15 minutos. **(REALIZAR)**
781. Por lo tanto, muchas mujeres quienes ya sufrían unas condiciones precarias perderán \*su empleo\*. **(PERDER)**
782. Excluyendo \*casos de la disputada República de Crimea\*, que fue anexionada por Rusia en 2014 pero sigue siendo reconocida internacionalmente como soberana de Ucrania. **(EXCLUIR)**
783. Si la DSMB encuentra que el tratamiento experimental no está funcionando o está perjudicando a los participantes, la Junta recomendará \*que el estudio se detenga de inmediato\*. **(RECOMENDAR)**
784. \*Los signos de coronavirus más frecuentes\* parecen haber cambiado. **(CAMBIAR)**
785. Los desechos deben ponerse \*en una bolsa plástica\*, cerrarla y botarla. **(PONER)**
786. Atendiendo a los datos del IMSERSO, las mujeres representan \*el 70,4% de las personas que viven en residencias\*, por un 29,6% de hombres. **(REPRESENTAR)**
787. Para ese entonces ya se habían comunicado \*los primeros casos por COVID-19 fuera de China: dos en Tailandia y uno en Japón\*. **(COMUNICAR)**
788. Todavía no se sabe con exactitud cómo o cuándo \*el virus\* se volvió infeccioso para las personas, pero sí se sabe que se transmite entre humanos. **(VOLVER)**
789. Sin embargo, \*la ausencia de anticuerpos en la sangre\* no significa que el sistema inmunitario no pueda producir nuevos anticuerpos rápidamente en caso de reexposición al SARS-CoV-2. **(SIGNIFICAR)**

790. El 31 de diciembre, el Comité de Salud Municipal de Wuhan informó a la Organización Mundial de la Salud (OMS) que \*27 personas\* habían sido diagnosticadas con neumonía de causa desconocida, habiendo 7 en estado crítico; la mayoría de estos casos eran trabajadores del mencionado mercado. **(DIAGNOSTICAR)**
791. Si la enfermedad no se trata a tiempo, \*estos síntomas\* pueden agravarse. **(AGRAVAR)**
792. \*Esto\* causa irritación y por ello empezamos a toser, señala Schaffner. **(CAUSAR)**
793. Los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) de los Estados Unidos y la Organización Mundial de la Salud (OMS) consideran \*que la COVID-19 es una grave amenaza para la salud pública\*. **(CONSIDERAR)**
794. Albert sale de la UCI y pide \*a todos\* que se vacunen **(PEDIR)**
795. Sin embargo, \*ahora\* se están realizando estudios clínicos para probar la seguridad y eficacia de las vacunas disponibles actualmente en niños más pequeños. **(REALIZAR)**
796. La FDA analizó \*los datos de los estudios clínicos\* de inmediato. **(ANALIZAR)**
797. También es esencial que usen su antebrazo o un pañuelo. (que deben desechar) cuando tosen o estornudan y que tras hacerlo siempre se laven \*las manos\* para evitar la propagación de otros virus, advierte. **(LAVAR)**
798. Tras analizar miles de casos de pacientes, los hospitales han determinado que existe \*un patrón de síntomas\* entre quienes se contagian de coronavirus. **(EXISTIR)**
799. \*Sanidad\* aprueba una cuarta dosis de la vacuna para personas inmunodeprimidas cinco meses después de la tercera. **(APROBAR)**
800. Sin embargo, las observaciones clínicas apuntan a que cuando la repuesta inmune no es capaz de controlar eficazmente el virus, como en personas mayores con un sistema inmune debilitado, el virus se propagaría de forma más eficaz produciendo daño en el tejido pulmonar, \*lo que\* activaría a los macrófagos y granulocitos y conduciría a la liberación masiva de citoquinas proinflamatorias. **(ACTIVAR)**
801. Al propagarse la pandemia por el hemisferio norte, \*la epidemia anual del gripe\* terminó abruptamente. **(TERMINAR)**
802. Albert sale de la UCI y pide a todos \*que se vacunen\* **(PEDIR)**
803. \*A comienzos de 2020\*, el mundo dirigió la mirada hacia China, donde el SARS-CoV-2 avanzaba por todo el país. **(AVANZAR)**
804. \*Un par de meses después\*, Europa se convertía en el foco de la pandemia, y en mayo, América era el continente más afectado del mundo. **(CONVERTIR)**
805. Hacerlo frecuentemente reduce hasta en 50% \*el riesgo de contraer coronavirus\*. **(REDUCIR)**
806. \*Italia\* mantiene la alerta en apertura de tráfico aéreo 3:33 **(MANTENER)**
807. Sin embargo, los investigadores no saben con qué frecuencia puede suceder \*esto\*. **(SUCEDER)**
808. Desecha la mascarilla usada y lávate \*las manos\*. **(LAVAR)**
809. Mantener al menos un metro de distancia de otras personas, particularmente \*aquellas que\* tosan, estornuden y tengan fiebre. **(ESTORNUDAR)**
810. \*Para el día de hoy\* no se han reportado defunciones nuevas. **(REPORTAR)**
811. La pandemia de la COVID-19 está en su segunda ola y ya ha dejado en nuestro país \*más de 900.000 personas infectadas\*. **(DEJAR)**
812. \*La prueba molecular\* puede tardar de unos días a una semana en proporcionar los resultados. **(TARDAR)**
813. El dolor de cabeza de la COVID suele tender a empeorar si la persona hace \*esfuerzos o movimientos de cabeza\*. **(HACER)**
814. Debido a la complejidad y coste de las pruebas PCR, \*en los primeros meses de la pandemia\* casi ningún país dispuso de datos confiables sobre la prevalencia del virus en su población. **(DISPONER)**
815. \*El gobierno de Hong Kong\* advirtió a cualquiera que viaje fuera de la ciudad que no toque animales; no coma carne de caza; y evite visitar mercados húmedos, mercados de aves vivas y granjas. **(ADVERTIR)**

816. La Organización Mundial de la Salud recomienda por este motivo que se desinfecten las superficies, especialmente \*en el entorno sanitario\*. **(DESINFECTAR)**
817. \*Esto\* se está investigando de cerca y se están realizando más estudios. **(INVESTIGAR)**
818. Después, desecha \*el pañuelo descartable\*, o lávalo si es de tela. **(DESECHAR)**
819. \*Otros síntomas más\* pueden parecerse a un resfriado o alergias, como dolor de garganta y congestión. **(PARECER)**
820. De todos modos, \*la duración y la gravedad de los síntomas\* depende de si la persona está vacunada o no. **(DEPENDER)**
821. Si los síntomas se incrementan, es adulto mayor o persona con comorbilidades, como diabetes o cáncer, debe acudir \*a la unidad de salud más cercana\* para recibir atención. **(ACUDIR)**
822. Por otro lado, los resultados de un estudio publicado en 'The Lancet Microbe', parecen indicar que \*la mayoría de la transmisión\* ocurre durante los primeros días, especialmente en los 5 primeros después del inicio de los síntomas, lo que indica la importancia del confinamiento después del inicio de los síntomas. **(OCURRIR)**
823. Unos investigadores han creado un algoritmo que permite a los relojes inteligentes \*tener la capacidad de detectar si tiene síntomas de la Covid-19\*. **(PERMITIR)**
824. Esto significa que la enfermedad se está extendiendo \*en varios países de todo el mundo\*, al mismo tiempo afectando a cientos de miles de personas en todos los continentes, causando un número de muertes que, a la fecha, sigue en ascenso. **(EXTENDER)**
825. \*Los investigadores de los NIH\* también están probando regímenes de vacunas mixtas. **(PROBAR)**
826. En el mes de mayo de 2021, otros 60 países han copatrocinado la propuesta, que alcanza el apoyo de más de 100 países, y el 5 de mayo los Estados Unidos dieron la sorpresa al anunciar que apoyaba \*la propuesta de suspensión de patentes\*, si bien solo en relación con la vacunas. **(APOYAR)**
827. Los datos de este estudio reflejan que apenas presentan esta característica el 13,9% de los afectados, pero sí puede ser \*un indicador que\* advierta la presencia del virus en nuestro organismo. **(ADVERTIR)**
828. Las infecciones asintomáticas son las más comunes entre niños, pero las afecciones dependen de la edad, ya que el dolor estómago es un síntoma común en los más pequeños, mientras que los mayores de 12 años pueden sufrir \*una pérdida del gusto o del olfato\*, según reconocieron un grupo de expertos en epidemiología. **(SUFRIR)**
829. \*El menor de 15 años\* manifestó que sí se quiere vacunar alegando que si se suministra es porque es confiable. **(MANIFESTAR)**
830. Más tarde, uno de los delegados declaró que el equipo quería \*llamarlo un patógeno peligroso\*, y sentían que la palabra peligroso podría estar vinculada al bioterrorismo. **(QUERER)**
831. El tiempo hace la diferencia, por ello es necesario que revise \*tus síntomas\* y acudas al hospital de acuerdo al semáforo de atención. **(REVISAR)**
832. \*Otra investigación\*, publicada en Frontiers in Public Health, también evidenció que el orden de las primeras afecciones sería: fiebre, tos seca y persistente y síntomas gastrointestinales –entre las que se encontraban las náuseas y la diarrea–. **(EVIDENCIAR)**
833. Paralelamente, cabe destacar la elevada presencia de mujeres en el ámbito sanitario, lo que ha supuesto \*que aproximadamente el 75% de todo el personal sanitario contagiado en España sean mujeres\*. **(SUPONER)**
834. Los investigadores han probado el nuevo algoritmo, desarrollado con software de código abierto, en 2.155 usuarios que \*durante ocho meses\* llevaron relojes inteligentes capaces de medir el estrés y las alteraciones en la frecuencia cardíaca, que si es más alta de lo normal puede ser señal clara de que existe algún problema. **(LLEVAR)**
835. Con mayor frecuencia, \*los síntomas\* aparecen alrededor de 5 días después de la exposición. **(APARECER)**
836. Hay mascarillas que llevan \*alguna válvula de exhalación (es decir, de salida\*), la cual expulsa el aire de la respiración sin filtrar. **(LLEVAR)**

837. Este se adapta durante el embarazo para permitir \*el desarrollo de un feto cuya carga genética solo comparte a medias con su madre\*, lo que conlleva a una reacción inmunitaria diferente frente a las infecciones durante la etapa del embarazo. **(PERMITIR)**
838. Se permiten visitas y acompañamiento a pacientes NO COVID \*en centros hospitalarios\*, bajo determinadas condiciones**(PERMITIR)**
839. Limpie el área u objeto \*con agua y jabón u otro detergente\* si está sucio. **(LIMPIAR)**
840. Pon los guantes y las máscaras usados \*en un bote de basura\* con tapa en el cuarto de la persona enferma. **(PONER)**
841. Además, la fiebre, tos seca y la pérdida del olfato y gusto, ya no son indicios tan comunes como al inicio de la pandemia, por ello esta lista de síntomas debe de cambiar, según manifiesta \*Tim Spector, líder del proyecto y profesor de epidemiología de la Universidad King's College de Londres\*. **(MANIFESTAR)**
842. Un informe de la Organización Mundial de la Salud incluye el dolor de garganta entre \*los síntomas que\* pueden padecer los pacientes con coronavirus. **(PADECER)**
843. Es importante recalcar que, si una mujer está en proceso de búsqueda del embarazo, la vacuna no afectará \*a su fertilidad en ningún caso\*. **(AFECTAR)**
844. El 14 de febrero murió \*un turista chino\* en París, en la que se creyó se trataba de la primera muerte por COVID-19 fuera de Asia pero posteriormente se descubrió que un hombre fallecido en la ciudad española de Valencia el 13 de febrero también estaba infectado de coronavirus, que había contraído en Nepal. **(MORIR)**
845. Dermatologic Therapy, \*en diciembre de 2020\*, reportó que algunos pacientes presentaban cambios en la apariencia de su lengua. **(REPORTAR)**
846. Los CDC han actualizado \*las recomendaciones para el público sobre el aislamiento y la cuarentena\* y están revisando el sitio web de los CDC para incluir estos cambios. **(ACTUALIZAR)**
847. Esto se debe sobre todo a \*la dependencia económica y a las dificultades para denunciar\* que han provocado e incrementado las nuevas medidas de confinamiento y aislamiento social. **(INCREMENTAR)**
848. Conforme a los hallazgos de la investigación británica, 1 de cada 5 pacientes experimenta también dolores \*en la zona abdominal\*. **(EXPERIMENTAR)**
849. Hay pacientes que son dados de alta, pero eso no significa \*que sus síntomas desaparecieron por completo\*. **(SIGNIFICAR)**
850. La FDA ha aprobado o autorizado \*varios tratamientos para enfermedades leves, moderadas y graves causadas por el COVID-19\*. **(APROBAR)**
851. La DSMB está compuesta por \*expertos en la afección específica del estudio que\* periódicamente analizan los resultados del estudio a medida que avanza. **(ANALIZAR)**
852. Deseche \*el pañuelo\* después de usarlo. **(DESECHAR)**
853. Sanidad aprueba una cuarta dosis de la vacuna \*para personas inmunodeprimidas\* cinco meses después de la tercera. **(APROBAR)**
854. La RTHK respondió revisando \*el programa\* y no encontró ninguna violación de su carta. **(REVISAR)**
855. Los pacientes que pierden el olfato no \*siempre\* pierden el gusto. **(PERDER)**
856. Italia mantiene \*la alerta\* en apertura de tráfico aéreo 3:33**(MANTENER)**
857. \*Los resultados de los datos de secuenciación\* se obtienen de la base de datos de GISAID. **(OBTENER)**
858. \*La Consejería de Sanidad\* recuerda que hay que seguir manteniendo las medidas habituales de uso de mascarillas en interiores y en el exterior en lugares concurridos, ventilación de interiores, distancia de seguridad e higiene de manos. **(RECORDAR)**
859. Para \*los pacientes hospitalizados con hipoxemia que\* requieren un suplemento de oxígeno convencional, el Panel recomienda usar una de las siguientes opciones: remdesivir (BIIa),

- dexametasona más remdesivir (BIII) o dexametasona sola (BI) (ver Manejo terapéutico de adultos con COVID-19). **(REQUERIR)**
860. Tras analizar miles de casos de pacientes, durante el brote en China a principios de este año, \*los hospitales\* han determinado que existe un patrón de síntomas entre quienes se contagian de coronavirus. **(DETERMINAR)**
861. Los estudios y los datos actuales muestran que \*los anticuerpos que\* nuestros cuerpos producen después de la vacunación reconocen y protegen contra las variantes del COVID-19. **(PRODUCIR)**
862. \*También nos da un formato estándar para usar en futuros brotes de otras y nuevas versiones del coronavirus\*, agregó. **(AGREGAR)**
863. La situación puede empeorar si el virus deja \*el conducto bronquial\* y llega a los pulmones, donde causa una inflamación (neumonía). **(DEJAR)**
864. \*El Instituto Feinstein de Northwell Health\* anunció en marzo un estudio sobre un anticuerpo humano que puede prevenir la actividad de IL-6. **(ANUNCIAR)**
865. Independientemente de las multas que pueden imponer \*las diferentes administraciones\*, los hechos podrían ser delito. **(IMPONER)**
866. En Estados Unidos, ya se está probando \*otra posible vacuna en humanos\*. **(PROBAR)**
867. \*El Nuevo Coronavirus COVID-19 es una cepa de la familia de coronavirus que\* no se había identificado previamente en humanos. **(IDENTIFICAR)**
868. Una vez que las copias están listas, salen de la célula \*donde se originaron\*, la destruyen y comienzan a infectar a otras células. **(SALIR)**
869. La Organización Mundial de la Salud recomienda por este motivo que se desinfecten \*las superficies\*, especialmente en el entorno sanitario. **(DESINFECTAR)**
870. Las precauciones acabarían \*cuando los peligros de la pandemia estén oficialmente terminados\*. **(ACABAR)**
871. Isaac González, publicista, recuerda así \*la vida que vuelve y la que se va, especialmente la de los ancianos, los más golpeados por la covid\* **(RECORDAR)**
872. La OMS desaconsejaba en marzo la utilización de máscara quirúrgica por la población sana, en abril \*la OMS\* consideró que era una medida aceptable en algunos países. **(CONSIDERAR)**
873. \*Lo que compartimos en el ardor del momento\* depende de nosotros. **(DEPENDER)**
874. La pandemia de COVID-19 en América es una epidemia continental que inició con la detección del primer caso de esta enfermedad el 21 de enero de 2020 en los Estados Unidos, \*un hombre de aproximadamente 30 años, originario del estado de Washington, que\* había viajado recientemente a China. **(VIAJAR)**
875. \*Álvarez, que es miembro del Comité Asesor Epidemiológico del Ministerio de Salud\*, explicó que los síntomas más comunes de la variante ómicron son los asociados con síntomas respiratorios altos, especialmente dolor de cabeza, producción de mocos, dolor de garganta, tos seca. **(EXPLICAR)**
876. El menor de 15 años manifestó \*que sí se quiere vacunar alegando que si se suministra es porque es confiable\*. **(MANIFESTAR)**
877. El virólogo y experto en influenza Robert Webster escribió \*en su libro Flu Hunter Unlocking the secrets of a virus (2019)\* que una pandemia mortal y disruptiva era solo cuestión de tiempo, y que la naturaleza eventualmente va a desafiar de nuevo a la humanidad con un equivalente al virus de la influenza de 1918. **(ESCRIBIR)**
878. Sin embargo, para la variante ómicron, la ausencia de esta afección y la presencia de las anteriores podría ser un indicio de que \*el paciente\* se ha contagiado con la nueva variante. **(CONTAGIAR)**
879. Sin embargo, la Food and Drug Administration (FDA) recomienda que las mujeres embarazadas o lactantes consulten con sus respectivos médicos \*a la hora de decidir si deben vacunarse\*. **(CONSULTAR)**
880. \*Las máscaras quirúrgicas (normales)\* son utilizadas desde hace tiempo por personas sanas en Hong Kong, Japón, Singapur y Malasia. **(UTILIZAR)**

881. Afirma que es imposible poner en un balanza los efectos adversos \*a medio y largo plazo\*. **(PONER)**
882. \*La vacunación generalizada\* ayudará a reducir el aumento de variantes adicionales. **(AYUDAR)**
883. La OMS está estudiando \*las investigaciones en curso sobre las formas de propagación, como las que apuntan a la transmisión aérea\*. **(ESTUDIAR)**
884. UNICEF está trabajando a contrarreloj \*con expertos en salud de todo el mund\* para proporcionar información rigurosa. **(TRABAJAR)**
885. La gripe suele ser la única que nos hace sentir \*dolor muscular\*. **(SENTIR)**
886. \*El coronavirus\* provoca pérdida de gusto y olfato, diarrea...**(PROVOCAR)**
887. \*Pacientes recientemente hospitalizados que han sido ingresados en la unidad de cuidados intensivos (UCI) dentro de las 24 horas anteriores y que\* requieren ventilación mecánica invasiva, ventilación mecánica no invasiva (VNI), u oxígeno a alto flujo por cánula nasal de (HFNC) > 0,4 FiO<sub>2</sub> / 30 L / min de flujo de oxígeno. (BIIa)**(REQUERIR)**
888. Cúbrase la boca y la nariz con un pañuelo desechable o su manga (no sus manos) \*cuando tosa o estornude\*. **(CUBRIR)**
889. La información de la situación epidemiológica y la evolución de la infección por coronavirus se actualiza con datos oficiales aportados por los diferentes sistemas de salud autonómicos al Centro de Coordinación de Alertas y Emergencias Sanitarias (CCAES), \*donde\* se pueden consultar los datos de España, Europa y el resto del Mundo. **(CONSULTAR)**
890. Sin embargo, la mayoría de los niños y adultos no serán considerados febriles hasta que su temperatura alcance \*los 100 grados Fahrenheit (37,7 grados Celsius)\*. **(ALCANZAR)**
891. Las rutas de transmisión de persona a persona del agente etiológico SARS-CoV-2 incluyen la transmisión directa por inhalación de microgotas y aerosoles liberadas a través de tos, estornudos, la respiración, el habla, los gritos o el canto,<sup>2</sup> o por contacto de las manos con superficies contaminadas, que luego tocan \*las membranas mucosas orales, nasales u oculares\*. **(TOCAR)**
892. \*El Gobierno vasco\* impone a la productora musical Last Tour la máxima sanción por no exigir el pasaporte covid en los recitales y permitir que nadie llevara protección**(IMPONER)**
893. \*El 23 de agosto de 2021\*, la FDA aprobó la vacuna Pfizer-BioNTech, llamada Comirnaty, para personas mayores de 16 años. **(APROBAR)**
894. Dermatologic Therapy, en diciembre de 2020, reportó \*que algunos pacientes presentaban cambios en la apariencia de su lengua\*. **(REPORTAR)**
895. \*La OMS\* está estudiando las investigaciones en curso sobre las formas de propagación, como las que apuntan a la transmisión aérea. **(ESTUDIAR)**
896. Pero ahora, la tendencia cambió y la tos ha sido superada \*por dolor de cabeza (66%), dolor de garganta (53%), secreción nasal (49%) y fiebre (42%)\* en menores de 40 años, mientras que los mayores de 40 mostraron resultados similares, precisó el profesor Spector. **(SUPERAR)**
897. Los expertos esperan \*que las vacunas estén autorizadas para niños y adolescentes a fines de 2021\*. **(ESPERAR)**
898. Además, los factores inflamatorios asociados con \*las enfermedades que\* contienen principalmente IL-6 aumentaron significativamente, lo que también contribuyó al agravamiento de la enfermedad alrededor de 2 a 10 días después del inicio. **(CONTENER)**
899. Hay pacientes que han tenido síntomas leves o, incluso, se han contagiado con el virus de forma asintomática, y los hay que han padecido una fuerte neumonía, dificultades para respirar o \*que\*, lamentablemente, han fallecido. **(FALLECER)**
900. Los estudios y los datos actuales muestran que los anticuerpos que nuestros cuerpos producen \*después de la vacunación\* reconocen y protegen contra las variantes del COVID-19. **(PRODUCIR)**
901. Actualmente, los estudios están inscribiendo a personas que están hospitalizadas con COVID-19 para probar la efectividad de varios medicamentos diferentes que se encuentran \*en una etapa avanzada de desarrollo clínico\*. **(ENCONTRAR)**

902. Incluso en el mejor de los casos, desde Naciones Unidas se afirma que \*las 1300 millones de personas del continente\* necesitarán 74 millones de equipos de prueba y 30 000 respiradores en 2020. **(NECESITAR)**
903. Si piensas \*que tienes síntomas de COVID-19\*, llama al médico. **(PENSAR)**
904. Esta crisis sanitaria ha evidenciado \*algunos problemas en materia de igualdad de género que existen en este país, sobre todo en el ámbito laboral y en especial en el sector de cuidados\*. **(EVIDENCIAR)**
905. En general, el coronavirus afecta menos a la parte alta del aparato respiratorio, a diferencia del resfriado común aunque, como decimos, \*las nuevas variantes\* han cambiado este patrón. **(CAMBIAR)**
906. Aún no hemos llegado \*a ese punto\*, ya que todavía tenemos que probarlo en ensayos clínicos, pero es el objetivo final, afirmó Michael Snyder. **(LLEGAR)**
907. Al 3 de marzo de 2020 a nivel global, \*3 110 de los 90 892 casos comunicados de COVID-19\* habían muerto. (3,4%), según cifras indicadas por el director de la OMS. **(MORIR)**
908. \*La investigación clínica\* encontró que se detecta una alta concentración de citoquinas en el plasma de pacientes críticos infectados con SARS-CoV-2, lo que sugiere que la tormenta de citocinas se asociaba con la gravedad de la enfermedad. **(ENCONTRAR)**
909. \*Una de cada seis personas que contrae el virus\* desarrolla una enfermedad grave y tiene gran dificultad para respirar, especialmente personas mayores y las que padecen afecciones médicas subyacentes, como hipertensión arterial, problemas cardíacos o diabetes. **(DESARROLLAR)**
910. Se administraron 600 mg de hidroxiclороquina (marca Plaquenil) \*a estos pacientes\* todos los días durante 10 días. **(ADMINISTRAR)**
911. Un par de meses después, \*Europa\* se convertía en el foco de la pandemia, y en mayo, América era el continente más afectado del mundo. **(CONVERTIR)**
912. Es importante recalcar que, si una mujer está en proceso de búsqueda del embarazo, \*la vacuna\* no afectará a su fertilidad en ningún caso. **(AFECTAR)**
913. El cuidador debe usar \*una mascarilla facial\* cuando esté cerca de la persona enferma, así como guantes si entra en contacto con su sangre, heces o fluidos corporales. **(USAR)**
914. El aislamiento se utiliza para separar \*a las personas que tienen el virus de la COVID-19\* de las personas que no están enfermas. **(SEPARAR)**
915. \*Esto\* significa cubrirte la boca y la nariz con el codo o pañuelo doblado cuando toses o estornudas. **(SIGNIFICAR)**
916. Está muriendo \*mucho gente\*. **(MORIR)**
917. El objetivo de la cuarentena es lograr \*que, en el caso de que la persona se hubiera infectado, no transmita a su vez la infección a otros\*. **(LOGRAR)**
918. Actualmente, los estudios están inscribiendo a personas que están hospitalizadas con COVID-19 para probar la efectividad de \*varios medicamentos diferentes que\* se encuentran en una etapa avanzada de desarrollo clínico. **(ENCONTRAR)**
919. Se recupera \*la actividad de baile en discotecas y establecimientos de ocio nocturno, con mascarilla obligatoria\*. **(RECUPERAR)**
920. \*El aumento de la utilización de internet\* ha causado un incremento de la violencia digital, y en especial de la violencia de género digital. **(CAUSAR)**
921. Además se ha recomendado evitar el contacto con animales, tanto vivos como muertos, en las zonas donde se está propagando el virus y, una recomendación para reducir también el riesgo de otras enfermedades infecciosas, solamente comer \*carne o huevos\* si están bien asados o cocinados. **(COMER)**
922. \*El profesor Tim Spector, epidemiólogo del King's College de Londres y que\* dirige el llamado Estudio de Síntomas Zoe Covid, explica que las personas más jóvenes que contraen la variante delta pueden sentirse más como si tuvieran un resfriado fuerte. **(DIRIGIR)**
923. \*IMSS\* aplicará prueba rápida de Covid-19 **(APLICAR)**

924. \*Los institutos y centros de los NIH, en asociación con empresas privadas\*, continúan estudiando diferentes tipos de vacunas. **(ESTUDIAR)**
925. Desde el día 2 de marzo se incluye \*información sobre la evolución de la vacunación\*; los datos de vacunación no se actualizan los fines de semana y festivos. **(INCLUIR)**
926. Los análisis preliminares también apuntan \*a que provoca una enfermedad bastante más leve\*, pero todavía es pronto para extraer conclusiones sobre ómicron. **(APUNTAR)**
927. Las personas con síntomas de resfriado deberían hacerse la prueba diagnóstica para frenar \*cualquier posible propagación del COVID\*. **(FRENAR)**
928. La idea es que las personas \*eventualmente\* utilicen esta información para decidir si necesitan hacerse una prueba de Covid-19 o aislarse. **(UTILIZAR)**
929. \*La fiebre y la tos\* se asocian frecuentemente con una variedad de enfermedades respiratorias, incluyendo el Síndrome Respiratorio del Medio Oriente (MERS) y el Síndrome Respiratorio Agudo Severo. (SARS). **(ASOCIAR)**
930. Así, por ejemplo, \*la iniciativa Global 50/50\* recoge en su página web información de 55 países, de los cuales únicamente 26 desglosan por sexo los datos actualizados de casos confirmados y de muertes. **(RECOGER)**
931. Sin embargo, con ómicron, la mialgia podría considerarse \*uno de los primeros indicios de que los pacientes se habían contagiado con la nueva versión del virus\*. **(CONSIDERAR)**
932. Si no es posible ponerte una mascarilla, cúbrete la boca y la nariz con un pañuelo o con el codo \*cuando tosas o estornudes\*. **(CUBRIR)**
933. Saber cómo actuar nos puede ayudar a controlar mejor \*las situaciones que\* puedan surgir estos días y a prestar ayuda a las personas de nuestro entorno. **(SURGIR)**
934. Al igual que para las salas de cine, se prevé que \*muchos espacios de exhibición\* no puedan superar un largo periodo de inactividad. **(SUPERAR)**
935. Aquí encontrará \*información actualizada\*. **(ENCONTRAR)**
936. Estos efectos secundarios raros generalmente aparecen \*dentro de las seis semanas después de recibir la dosis de la vacuna\*. **(APARECER)**
937. Una vez que \*las copias\* están listas, salen de la célula donde se originaron, la destruyen y comienzan a infectar a otras células. **(SALIR)**
938. Si la situación del paciente sigue agravándose, este es el momento en el que es probable \*que\* ingresen a la unidad de cuidados intensivos (UCI). **(INGRESAR)**
939. Si el COVID-19 es tratable, ¿necesito \*una vacuna\*? **(NECESITAR)**
940. Aunque las personas de avanzada edad son más vulnerables a la enfermedad, a mitad de marzo de 2020, la OMS recordó \*que los jóvenes también pueden tener complicaciones e incluso llegar a la muerte por el virus\*. **(RECORDAR)**
941. \*A veces\* el primer caso de COVID-19 identificado por el sistema sanitario. (caso índice o paciente 0), no coincide con el caso que da origen a la epidemia. **(COINCIDIR)**
942. \*El virus que\* salió de Wuhan no es exactamente el mismo que entró en España. **(SALIR)**
943. Para el 6 de enero de 2022, se habían administrado \*9 370 millones de dosis de las vacunas contra la COVID-19\*, distribuidas solamente entre el 59% de la población mundial. **(ADMINISTRAR)**
944. Lavarse las manos con frecuencia durante 20 segundos \*con agua y jabón\* es fundamental para el cuidador. **(LAVAR)**
945. En cambio, toma su temperatura \*al final de la tarde y al anochecer\*. **(TOMAR)**
946. En cambio, ha incrementado \*el número de casos de dolor de cabeza, secreción nasal, estornudos y dolor de garganta\*. **(INCREMENTAR)**
947. \*Para el 1 de enero de 2020\*, el mercado había sido cerrado y se había descartado que el causante de la neumonía fuera el SARS, el MERS, gripe, gripe aviaria u otras enfermedades respiratorias comunes causadas por virus. **(CERRAR)**
948. En cambio, \*los respiradores certificados\* también dan una gran protección a su portador. **(DAR)**
949. Lo que comenzó como un brote local en la provincia china de Hubei se convirtió en menos de tres meses \*en una pandemia global que afecta ya a casi todo el planeta\*. **(CONVERTIR)**

950. A pesar de que las autoridades del Servicio Nacional de Salud inglés (NHS) reconocen a la temperatura alta, tos persistente y a la pérdida o cambio repentino en el sentido de olfato o gusto. (anosmia y/o disgeusia) como principales síntomas del coronavirus, \*investigadores del King's College de Londres que\* dirigen el proyecto de vigilancia más ambicioso del mundo denominado COVID ZOE Symptom Study alertaron que la enfermedad pandémica está ahora actuando de manera diferente. (DIRIGIR)
951. Algunos pacientes permanecen \*en el hospital\* por más tiempo. (PERMANECER)
952. Sin embargo, Michael Snyder, uno de los autores del estudio, aseguró que el hallazgo más importante fue que pudieron detectar el 80% de los casos confirmados de coronavirus \*antes o cuando los participantes presentaban síntomas\*. (DETECTAR)
953. España es el segundo por número de contagios, aunque ya empieza a ver \*la luz\* al final del túnel. (VER)
954. Vivimos \*un momento difícil\* y tener a los niños en casa, puede convertir nuestra rutina remota en una actividad muy particular. (VIVIR)
955. \*Esta terapia\* fue autorizada por la FDA estadounidense en agosto de 2020 como tratamiento de emergencia contra la COVID-19. (AUTORIZAR)
956. Las Naciones Unidas están siguiendo de cerca \*la situación del brote de COVID-19\*. (SEGUIR)
957. El Gobierno vasco impone a la productora musical Last Tour \*la máxima sanción\* por no exigir el pasaporte covid en los recitales y permitir que nadie llevara protección (IMPONER)
958. Son varios los síntomas de la covid-19, pero \*hasta ahora\* no se ha conocido la ruta más probable de sintomatología. (CONOCER)
959. Además, \*las circunstancias actuales\* dificultan el acceso a teléfonos, líneas de atención y servicios públicos como la policía, la justicia y los servicios sociales, por lo que elevan la vulnerabilidad de las mujeres a la violencia en el ámbito privado en todo el mundo. (ELEVAR)
960. El Instituto Feinstein de Northwell Health anunció \*en marzo\* un estudio sobre un anticuerpo humano que puede prevenir la actividad de IL-6. (ANUNCIAR)
961. La vacunación generalizada ayudará \*a reducir el aumento de variantes adicionales\*. (AYUDAR)
962. La persona encargada de los cuidados u otra persona debería usar \*una mascarilla y guantes desechables\* antes de limpiar. (USAR)
963. De acuerdo a la explicación de la OMS, el nuevo nombre se toma de las palabras corona, virus y disease (enfermedad en inglés), mientras que 19 representa el año en que surgió el brote se informó \*a la OMS\* el 31 de diciembre de 2019. (INFORMAR)
964. Así que los vacunados todavía deben realizar resoluciones preventivas (lavarse las manos con jabón, ponerse mascarilla, mantener la distancia social, etc.) para que la enfermedad no se propague, especialmente \*hacia gente vulnerable\*. (PROPAGAR)
965. Uno de los misterios que aún rodean a \*este coronavirus\* está en su capacidad de mutar, algo que será clave para entender si bastará con una vacuna para erradicarlo o si por el contrario volverá cada nueva temporada con una forma distinta. (VOLVER)
966. Realiza el lavado de manos frecuentemente o bien, desinfectarlas \*con gel con base de alcohol al 70%\*. (DESINFECTAR)
967. \*Todo el mundo\* habla del coronavirus. (HABLAR)
968. Cada virus puede crear \*entre 10.000 y 100.000 réplicas\*. (CREAR)
969. Además, los factores inflamatorios asociados con las enfermedades que contienen principalmente \*IL-6\* aumentaron significativamente, lo que también contribuyó al agravamiento de la enfermedad alrededor de 2 a 10 días después del inicio. (CONTENER)
970. Limpia y desinfecta \*las canastas para la ropa\* y lávate las manos después. (LIMPIAR)
971. Sin embargo, usted puede propagar \*el virus\* incluso cuando no tiene síntomas. (PROPAGAR)
972. Unos investigadores han creado \*un algoritmo que permite a los relojes inteligentes tener la capacidad de detectar si tiene síntomas de la Covid-19\* (CREAR)
973. El primer caso de la pandemia de COVID-19 en Asia y en el mundo se detectó en Wuhan y finalmente se extendió \*al resto de la República Popular China\*. (EXTENDER)

974. En Italia se cuentan como víctimas de la pandemia \*a los muertos que habían dado positivo\* pero no a muchos ancianos fallecidos en residencias sin haberles hecho la prueba; mientras que en Bélgica se contabiliza en las cifras de fallecidos a todos los muertos en residencias con síntomas similares a los de COVID-19, aunque no se les hayan hecho tests. **(CONTAR)**
975. Ir al médico en caso de fiebre, tos y dificultad para respirar, llamando con antelación si se encuentra en zonas donde se está propagando el virus o si se \*las\* han visitado en los últimos 14 días. **(VISITAR)**
976. El hospital informó \*esto\* al CCDC, cuyo equipo en la ciudad inició una investigación. **(INFORMAR)**
977. Otra investigación, publicada en *Frontiers in Public Health*, también evidenció \*que el orden de las primeras afecciones sería: fiebre, tos seca y persistente y síntomas gastrointestinales –entre las que se encontraban las náuseas y la diarrea–\*. **(EVIDENCIAR)**
978. En adición, en la parte de *¿Quiénes se han vacunado?* se puede seleccionar el rango de tiempo para el que se quieren \*estos datos\*. **(QUERER)**
979. La industria cinematográfica china había perdido US\$2 mil millones \*en marzo de 2020\*, después de haber cerrado todos sus cines durante el período del Año Nuevo Lunar que sustenta a la industria en toda Asia. **(PERDER)**
980. El espectro clínico de COVID-19 varía de formas asintomáticas a condiciones clínicas caracterizadas por insuficiencia respiratoria severa que requiere ventilación mecánica y apoyo \*en una unidad de cuidados intensivos (UCI)\*, a manifestaciones en varios órganos y sistemas produciendo sepsis, shock séptico y síndrome de disfunción multiorgánica. **(REQUERIR)**
981. Por el contrario, los síntomas más comunes en jóvenes vacunados son fiebre, dolor de cabeza, tos persistente y dolor de garganta, reconoció \*el profesor Duncan\*. **(RECONOCER)**
982. Las gotas en el aire también pueden caer \*sobre superficies y objetos\*. **(CAER)**
983. Quienes ya se estaban recuperando pueden ver que \*sus fiebres\* desaparecen después de 12 días. **(DESAPARECER)**
984. El coronavirus causante del covid-19 se ha cobrado la vida de unas 286.000 personas –26.920 de ellas en España– y ha infectado a más de 4.177.000 \*en todo el mundo\*, según el último balance global ofrecido por la Universidad Johns Hopkins. **(INFECTAR)**
985. Mantenerse informado y seguir \*las recomendaciones de los profesionales médicos\*. **(SEGUIR)**
986. Algunos tratamientos propuestos no han demostrado \*ser efectivos\* y podrían ser peligrosos. **(DEMOSTRAR)**
987. Se han descrito \*casos de enfermos de COVID que\* muestran problemas neurológicos de alteración del nivel de conciencia. **(MOSTRAR)**
988. La enfermedad es mortal en raras ocasiones, y hasta ahora las víctimas mortales han sido personas de edad avanzada que ya padecían \*una enfermedad crónica como diabetes, asma o hipertensión\*. **(PADECER)**
989. En cuanto a estado de vacunación, se muestra \*el número de personas aptas en Puerto Rico que\* han completado una serie de vacunas COVID-19. **(COMPLETAR)**
990. De igual modo, en Sudán del Sur, \*la situación\* se agravó considerablemente durante el confinamiento. **(AGRAVAR)**
991. De ello informa el proyecto Zoe, una iniciativa que cuenta con \*más de cuatro millones de voluntarios\* que analiza la evolución de los síntomas del coronavirus. **(CONTAR)**
992. El estudio describió \*un subconjunto único de casos más leves en los que los síntomas iniciales fueron problemas digestivos como diarrea, a menudo sin fiebre\*. **(DESCRIBIR)**
993. IMSS aplicará \*prueba rápida de Covid-19\* **(APLICAR)**
994. Corea del Norte sigue siendo el único país en el que aún se desconoce \*algún caso confirmado de COVID-19\*, aunque fuentes del gobierno de Corea del Sur sospechan que la enfermedad ha afectado al país. **(DESCONOCER)**
995. \*La situación\* puede empeorar si el virus deja el conducto bronquial y llega a los pulmones, donde causa una inflamación (neumonía). **(EMPEORAR)**

996. Este mapa dispone \*de un buscador\* (recuadro arriba a la izquierda) para localizar una calle o domicilio; el resultado sitúa la zona que corresponde. **(DISPONER)**
997. Dado que el SARS-CoV-2 ha estado en la población humana solo desde diciembre de 2019, al principio de la pandemia se desconocía \*si la inmunidad era duradera en las personas que se recuperan de la enfermedad\*. **(DESCONOCER)**
998. Durante las primeras olas de infección se hizo \*el primer gran estudio sobre los síntomas\*, y que publicó la prestigiosa The New England Journal of Medicine que acaloraba qué síntomas no eran tan habituales. **(HACER)**
999. Evite el contacto \*con mascotas u otros animales\*. **(EVITAR)**
1000. \*UNICEF\* está trabajando a contrarreloj con expertos en salud de todo el mundo para proporcionar información rigurosa. **(TRABAJAR)**
1001. Según la OMS, en situaciones similares con otros virus, las mujeres también han alcanzado \*un mayor índice de contagios\*. **(ALCANZAR)**
1002. La Casa Blanca anunció que todos los viajeros internacionales que ingresan \*a los Estados Unidos\* deberán estar vacunados y este requisito comenzará a regir a partir del 8 de noviembre de 2021. **(INGRESAR)**
1003. \*Los CDC\* han actualizado algunas de las formas de operar los sistemas de atención médica de manera efectiva en respuesta a la vacunación contra el COVID-19. **(ACTUALIZAR)**
1004. Se han identificado casos en todos los continentes y, el 6 de marzo se confirmó \*el primer caso\* en Colombia. **(CONFIRMAR)**
1005. También nos estamos enfocando en la participación comunitaria al movilizar nuestra amplia red social para generar \*confianza y aceptación de las vacunas\*. **(GENERAR)**
1006. Debe permanecer en casa y evitar el contacto con personas hasta que \*su proveedor\* le indique que es seguro terminar el aislamiento en casa. **(INDICAR)**
1007. El hospital informó esto \*al CCDC\*, cuyo equipo en la ciudad inició una investigación. **(INFORMAR)**
1008. \*Para el 6 de enero de 2022\*, se habían administrado 9 370 millones de dosis de las vacunas contra la COVID-19, distribuidas solamente entre el 59% de la población mundial. **(ADMINISTRAR)**
1009. El infectólogo Carlos Álvarez afirmó \*que ante síntomas gripales lo mejor es aislarse ya que podría ser COVID-19\*. **(AFIRMAR)**
1010. \*El informe final\* se refiere al coronavirus como un nuevo patógeno que es altamente contagioso, puede propagarse rápidamente y debe considerarse capaz de causar enormes impactos en la salud, la economía y la sociedad. **(REFERIR)**
1011. Por ejemplo, se ha observado que las organizaciones de medios estatales People's Daily y Global Times, junto con el subdirector de información Zhao Lijian del Ministerio de Relaciones Exteriores, publican elogios efusivos sobre la respuesta de Beijing a la epidemia, como una amplia cobertura de \*la construcción acelerada de los nuevos hospitales en Wuhan\* (que Zhao afirmó se completó en 16 horas), la cuarentena de Wuhan con su población de 11 millones y la cuarentena sin precedentes de la provincia de Hubei. **(COMPLETAR)**
1012. Aunque las personas de avanzada edad son más vulnerables a la enfermedad, \*a mitad de marzo de 2020\*, la OMS recordó que los jóvenes también pueden tener complicaciones e incluso llegar a la muerte por el virus. **(RECORDAR)**
1013. \*Una mujer china de 61 años vecina de Wuhan que\* el 8 de enero viajó en vuelo directo de Wuhan a Tailandia en compañía de cinco miembros de su familia en un grupo de dieciséis personas. **(VIAJAR)**
1014. Esos pacientes experimentaron demoras en las pruebas y el diagnóstico que los pacientes con problemas respiratorios, y tardaron más en eliminar el virus \*de sus sistemas\*. **(ELIMINAR)**
1015. Menos personas informaron \*que perdieron el sentido del gusto\*. **(INFORMAR)**
1016. A pesar de que las autoridades del Servicio Nacional de Salud inglés (NHS) reconocen a la temperatura alta, tos persistente y a la pérdida o cambio repentino en el sentido de olfato o gusto. (anosmia y/o disgeusia) como principales síntomas del coronavirus, investigadores del King's

- College de Londres que dirigen el proyecto de vigilancia más ambicioso del mundo denominado COVID ZOE Symptom Study alertaron que \*la enfermedad pandémica\* está ahora actuando de manera diferente. **(ACTUAR)**
1017. \*Las precauciones\* acabarían cuando los peligros de la pandemia estén oficialmente terminados. **(ACABAR)**
1018. \*FDA\* aprobó una vacuna contra el COVID-19 y dos fueron autorizadas para su uso de emergencia. **(APROBAR)**
1019. La de 'tipo L' presentaba especial virulencia y fue más común al comienzo del brote, pero \*su frecuencia\* ha disminuido después de principios de enero, a causa, según los científicos, de las estrictas medidas que China puso en marcha. **(DISMINUIR)**
1020. \*Los médicos\* pueden enviar a los pacientes leves a recuperarse en su casa, bajo las siguientes condiciones: **(ENVIAR)**
1021. Pese a que muchas personas han comparado \*este virus\* con el de la gripe, la OMS cifra en un 3,4% la proporción de casos que terminan con la muerte del paciente. **(COMPARAR)**
1022. Si bien cuenta con una alta sensibilidad y especificidad, \*el procesamiento de muestras y en consecuencia la emisión del resultado\* puede tardar varias horas. **(TARDAR)**
1023. \*El hospital\* informó esto al CCDC, cuyo equipo en la ciudad inició una investigación. **(INFORMAR)**
1024. La variante ómicron se ha notificado hasta ahora \*en 19 países y territorios de la región\*. **(NOTIFICAR)**
1025. Sin embargo, \*casi 1 de cada 5 personas\* con una infección por coronavirus se enfermará gravemente y desarrollará dificultades respiratorias. **(ENFERMAR)**
1026. El Sindicat d'Habitatge de Nou Barris cree que se están acumulando los desalojos en previsión de que acabe \*la moratoria\*. **(ACABAR)**
1027. La idea es perfeccionar las alertas para que puedan avisar de qué es lo que realmente está sucediendo y para ello planean agregar \*datos\* al algoritmo como los patrones de sueño, el recuento de pasos y la temperatura corporal. **(AGREGAR)**
1028. Según la Superintendencia de Salud, \*las Isapres\* al ser parte de la seguridad social y de la protección del derecho a la salud, no podrán, en ningún caso, excluir la cobertura a la enfermedad aunque haya sido declarada pandemia. **(EXCLUIR)**
1029. Y aunque \*los datos\* mejoran en el tercer y cuarto trimestre, sigue sin haber una disminución considerable en la violencia hacia las mujeres. **(MEJORAR)**
1030. Este tipo de epidemias suelen estar provocadas por \*virus de animales que\* usan ácido ribonucleico. (ARN) como material genético para mutar y acaban transmitiéndose a otros seres vivos. **(USAR)**
1031. El 18 de diciembre de 2020, \*la FDA\* autorizó el uso de emergencia de esta vacuna contra el COVID-19 financiada por los NIH en los Estados Unidos para personas mayores de 18 años. **(AUTORIZAR)**
1032. ORDEN 1244/2021, de 1 de octubre, de la Consejería de Sanidad, por la que se establecen \*medidas preventivas\* para hacer frente a la crisis sanitaria ocasionada por el COVID-19. **(ESTABLECER)**
1033. Al 3 de marzo de 2020 \*a nivel global\*, 3 110 de los 90 892 casos comunicados de COVID-19 habían muerto. (3,4%), según cifras indicadas por el director de la OMS. **(MORIR)**
1034. Los Centros para el Control y la Prevención de Enfermedades (CDC) no pueden dar fe sobre la precisión de \*un sitio web que\* no pertenece al Gobierno Federal. **(PERTENECER)**
1035. \*La fatiga\* puede continuar mucho después de que el virus desaparezca. **(CONTINUAR)**
1036. \*Cuando tú o tu ser querido se recuperen\*, busca apoyo emocional. **(BUSCAR)**
1037. Si una persona obtiene un resultado positivo en COVID-19, debe llamar al teléfono 900 102 112 para que \*le\* recojan sus datos y le informen de las medidas que debe adoptar. **(RECOGER)**
1038. Todo el mundo habla de la enfermedad por coronavirus 2019 (COVID-19), y allá donde mires hay \*información sobre el virus y sobre cómo protegerte\*. **(HABER)**

1039. La conjuntivitis, una afección altamente contagiosa causada por un virus, es una inflamación de \*la capa delgada y transparente de tejido, llamada conjuntiva, que\* cubre la parte blanca del ojo y el interior del párpado. **(CUBRIR)**
1040. \*Buena parte de las mujeres\* trabajan en el sector de la restauración y del turismo, de los más afectados por la pandemia, además de ser las principales cuidadoras de personas enfermas, de niños y de personas de la tercera edad. **(TRABAJAR)**
1041. Los científicos de los NIH han estado estudiando \*las vacunas de ARNm y los coronavirus\* durante décadas. **(ESTUDIAR)**
1042. Los investigadores han descubierto \*que el virus puede permanecer en el aire durante, al menos, 30 minutos y viajar hasta 4,5 metros\*. **(DESCUBRIR)**
1043. \*Los síntomas de la COVID-19\* pueden variar de leves a graves. **(VARIAR)**
1044. \*El gobierno federal\* está proporcionando vacunas contra el COVID-19 autorizadas por la FDA de forma gratuita a todas las personas que viven en los Estados Unidos, independientemente de su estado migratorio o de seguro médico. **(PROPORCIONAR)**
1045. Por otra parte, para octubre de 2020 y 9370 millones de personas vacunadas en segmento del 59% de la población mundial, la Organización Mundial de la Salud estimaba \*que al menos el 10 % de la población mundial ya se había contagiado de esta enfermedad (unas 780 millones de personas infectadas aproximadamente)\*, debido al gran subregistro de casos a nivel mundial. **(ESTIMAR)**
1046. Pero 125 niños, casi el 6 %, desarrollaron una enfermedad muy grave y \*un niño de 14 años con infección confirmada por coronavirus\* murió. **(MORIR)**
1047. \*Las infecciones asintomáticas son las más comunes entre niños, pero las afecciones dependen de la edad, ya que el dolor estómago es un síntoma común en los más pequeños, mientras que los mayores de 12 años pueden sufrir una pérdida del gusto o del olfato\*, según reconocieron un grupo de expertos en epidemiología. **(RECONOCER)**
1048. Los científicos de los NIH han estado estudiando las vacunas de ARNm y los coronavirus \*durante décadas\*. **(ESTUDIAR)**
1049. Llama al 911 o al número local para emergencias si no puedes despertar a la persona enferma o si notas \*signos que\* indiquen una emergencia. **(INDICAR)**
1050. Si dudas y no te han hecho \*el test\*, mejor extrema el aislamiento. **(HACER)**
1051. Al vacunarse, también puede proteger \*a sus seres queridos\*. **(PROTEGER)**
1052. \*Una prueba viral\* busca una infección actual. **(BUSCAR)**
1053. Es fundamental que la población no se relaje, no baje la guardia para que no perdamos \*lo que\* hemos logrado en los últimos meses: la reducción y estabilidad de la epidemia y su control asistencial. **(LOGRAR)**
1054. La Organización Mundial de la Salud recomienda por este motivo \*que se desinfecten las superficies\*, especialmente en el entorno sanitario. **(RECOMENDAR)**
1055. Así avanza la vacunación \*en España\*: cuatro de cada diez niños ya tienen la primera dosis de la vacuna. **(AVANZAR)**
1056. \*Después\*, desecha el pañuelo descartable, o lávalo si es de tela. **(DESECHAR)**
1057. Limpie y desinfecte \*las superficies de contacto frecuente\* en la habitación del enfermo y el baño; use guantes desechables. **(DESINFECTAR)**
1058. \*Algunas organizaciones internacionales, como la OMS\*, han publicado medidas preventivas para reducir la transmisión del virus. **(PUBLICAR)**
1059. \*La gripe\* pertenece a la familia de la ortomixovirus, en que la que también se ubican la influenza estacionaria que se presenta cada año y puede cobrar vidas por sí misma. **(PERTENECER)**
1060. Al igual que a la de otros virus, la respuesta inmunitaria a la COVID-19 dependerá de si se dispone \*de un sistema inmunitario sano\*. **(DISPONER)**
1061. Asimismo, se debe evitar el contacto estrecho \*con cualquier persona que presente signos de afección respiratoria\*, como tos o estornudos. **(EVITAR)**
1062. El 23 de agosto de 2021, la FDA aprobó \*la vacuna Pfizer-BioNTech, llamada Comirnaty\*, para personas mayores de 16 años. **(APROBAR)**

1063. \*Esta página\* compara el COVID-19 con la influenza, con base en la mejor información disponible a la fecha. **(COMPARAR)**
1064. Este tipo de epidemias suelen estar provocadas por virus de animales que usan \*ácido ribonucleico. (ARN)\* como material genético para mutar y acaban transmitiéndose a otros seres vivos. **(USAR)**
1065. Además, \*las personas\* pueden tardar más en presentar síntomas y ser contagiosas por más tiempo. **(TARDAR)**
1066. El primer caso de la pandemia de COVID-19 en Asia y en el mundo se detectó en Wuhan y \*finalmente\* se extendió al resto de la República Popular China. **(EXTENDER)**
1067. Use un vaporizador o tome \*una ducha de vapor\* para aumentar la humedad del aire y ayudar a calmar la garganta seca y la tos. **(TOMAR)**
1068. Esos pacientes experimentaron demoras en las pruebas y el diagnóstico que los pacientes con problemas respiratorios, y tardaron más en eliminar \*el virus\* de sus sistemas. **(ELIMINAR)**
1069. \*El nuevo Coronavirus\* causa una Infección Respiratoria Aguda (IRA), es decir una gripa, que puede ser leve, moderada o severa. **(CAUSAR)**
1070. La Resolución 380 de 2020 establece \*que tanto los viajeros nacionales como extranjeros provenientes de China, Italia, Francia y España deben adoptar medidas de aislamiento e internación en su residencia o en un hospedaje transitorio cubierto por su propia cuenta durante 14 días\*. **(ESTABLECER)**
1071. \*Varias empresas farmacéuticas, como Vir Biotechnology\*, han evaluado la efectividad contra el virus SARS-Cov-2 de anticuerpos monoclonales previamente identificados contra otros virus. **(EVALUAR)**
1072. Quédese \*en casa\*, excepto para conseguir atención médica. **(QUEDAR)**
1073. \*A veces\*, las personas que están completamente vacunadas contra una infección irruptiva, lo que significa que dan positivo en la prueba del SARS-CoV-2 o se enferman con COVID-19. **(CONTRAER)**
1074. \*El virus\* puede pasar a través de las mucosas, especialmente la mucosa nasal y laríngea, luego ingresa a los pulmones a través del tracto respiratorio. **(PASAR)**
1075. Si la DSMB encuentra que el tratamiento experimental no está funcionando o está perjudicando a los participantes, \*la Junta\* recomendará que el estudio se detenga de inmediato. **(RECOMENDAR)**
1076. Cámbiate \*la mascarilla\* todos los días. **(CAMBIAR)**
1077. Si sus síntomas no indican \*una emergencia médica\*, su proveedor puede decidir controlar sus síntomas mientras se recupera en casa. **(INDICAR)**
1078. Mantenga \*un registro de sus síntomas\*. **(MANTENER)**
1079. Una persona acatarrada tiene \*congestión nasal, estornudos\* y puede que fiebre baja y tos. **(TENER)**
1080. Los adultos mayores y aquellos con afecciones crónicas graves se enfrentan a un mayor riesgo de enfermedades graves y complicaciones por COVID-19 y los CDC de Estados Unidos les han aconsejado \*que eviten las multitudes y se queden en casa tanto como sea posible en áreas de brote comunitario\*. **(ACONSEJAR)**
1081. Evite \*el contacto\* con mascotas u otros animales. **(EVITAR)**
1082. El COVID-19 ha demostrado, \*desde las primeras olas de la pandemia\*, que las enfermedades físicas también pueden causar afecciones mentales. **(DEMOSTRAR)**
1083. \*La gente vacunada\* todavía puede transmitir la COVID, aunque es más difícil. **(TRANSMITIR)**
1084. Además, se sabe \*que hay otros coronavirus circulando entre animales, que todavía no han infectado al ser humano\*. **(SABER)**
1085. \*Signos que\* advierten que se trata de una emergencia. **(ADVERTIR)**
1086. Si \*los síntomas\* parecen estar empeorando, llama al médico. **(EMPEORAR)**
1087. Al principio, la ciencia no pensaba \*que la diarrea u otros problemas gástricos típicos que a menudo vienen con la gripe se aplicaran al nuevo coronavirus, también conocido como SARS-CoV-2\*. **(PENSAR)**

1088. Los científicos creen \*que la variante que nació en India y clasificada por la OMS como Delta, podría estar detrás del cambio en el cuadro clínico\*(**CREER**)
1089. Rápidamente, \*a principios de enero\*, las autoridades de este país identificaron la causa como una nueva cepa de coronavirus. (**IDENTIFICAR**)
1090. \*Un estudio de Wuhan\* descubrió que, en promedio, pasaron 12 días hasta que los pacientes fueron admitidos en la UCI. (**DESCUBRIR**)
1091. El primer caso de COVID-19 en Europa se detectó el 25 de enero en Francia: un hombre de 80 años procedente de la provincia de Hubei, que luego falleció \*el 15 de febrero\*. (**FALLECER**)
1092. En diciembre de 2019 hubo \*un brote epidémico de neumonía de causa desconocida\* en Wuhan, provincia de Hubei, China; el cual, según afirmó más tarde Reporteros Sin Fronteras, llegó a afectar a más de 60 personas el día 20 de ese mes. (**HABER**)
1093. Los observadores han advertido que si bien la admiración de los trabajadores médicos de primera línea es generalizada y sincera, \*los medios estatales\* también deberían destacar la realidad de que muchos de esos trabajadores carecen de equipo de protección y que más de 3000 han sido infectados desde el brote y que la atención de los medios puede brindarles apoyo público para obtener algunos equipos muy necesarios. (**DESTACAR**)
1094. El 31 de diciembre de 2019, la Organización Mundial de la Salud (OMS) recibió reportes de presencia de neumonía, de origen desconocido, \*en la ciudad de Wuhan, en China\*. (**RECIBIR**)
1095. Esta terapia fue autorizada por la FDA estadounidense \*en agosto de 2020\* como tratamiento de emergencia contra la COVID-19. (**AUTORIZAR**)
1096. Por otra parte, en los casos graves se ha observado \*un mecanismo inmunopatológico mediado por un aumento de IL-2, IL-7, IL-17, IL-10, MCP-1, MIP-1a y TNF- $\alpha$ \*. (**OBSERVAR**)
1097. El virus SARS-CoV-2 puede ingresar \*al cuerpo\* cuando otra persona inhala estas gotas. (**INGRESAR**)
1098. Conforme a las indicaciones de los Centros para el Control y Prevención de Enfermedades (CDC, por sus siglas en inglés), \*los pacientes\* podrían experimentar náuseas o vómitos. (**EXPERIMENTAR**)
1099. \*Menos personas\* informaron que perdieron el sentido del gusto. (**INFORMAR**)
1100. Agradecemos la iniciativa GISAID, los laboratorios responsables de obtener las muestras, así como los laboratorios que generaron \*las secuencias genómicas\* y compartieron a través de GISAID. (**GENERAR**)
1101. Existe mucha información errónea \*en línea\* sobre los tratamientos para el COVID-19. (**EXISTIR**)
1102. Un informe publicado por la Organización Mundial de la Salud en febrero descubrió que más del 33% de 55.924 personas con casos confirmados por laboratorio de covid-19 habían tosido \*esputo\*, un moco espeso a veces llamado flema, de sus pulmones. (**TOSER**)
1103. Se reportaron casos de reinfección en \*pacientes que\* ya habían pasado la enfermedad y se temía que el SARS-CoV-2 se comportase como algunos otros coronavirus, que son capaces de reinfectar después de aproximadamente un año. (**PASAR**)
1104. El Instituto Feinstein de Northwell Health anunció en marzo un estudio sobre un anticuerpo humano que puede prevenir \*la actividad de IL-6\*. (**PREVENIR**)
1105. La partícula viral saldrá \*de la célula\* por exocitosis. (**SALIR**)
1106. El 10 de enero de 2020, las primeras secuencias del genoma completo del SARS-CoV-2 estuvieron disponibles en GISAID, lo que permitió \*respuestas globales a la pandemia\*, incluyendo el desarrollo de las primeras vacunas y pruebas de diagnóstico para detectar el SARS-CoV-2. (**PERMITIR**)
1107. De ello informa el proyecto Zoe, \*una iniciativa que\* cuenta con más de cuatro millones de voluntarios que analiza la evolución de los síntomas del coronavirus. (**CONTAR**)
1108. Las máscaras también se recomiendan para quienes cuidan \*a alguien que pueda tener la enfermedad\*. (**CUIDAR**)
1109. Una persona con fiebre sentirá \*que su espalda o pecho están calientes\*. (**SENTIR**)

1110. Los datos pasan por un proceso que requiere que \*el personal del Departamento de Salud de Puerto Rico\* revise, importe y asigne casos. **(REVISAR)**
1111. Según la OMS, en situaciones similares con otros virus, \*las mujeres\* también han alcanzado un mayor índice de contagios. **(ALCANZAR)**
1112. El cuidador debe usar una mascarilla facial \*cuando esté cerca de la persona enferma\*, así como guantes si entra en contacto con su sangre, heces o fluidos corporales. **(USAR)**
1113. La conjuntivitis, una afección altamente contagiosa causada por un virus, es una inflamación de la capa delgada y transparente de tejido, llamada conjuntiva, que cubre \*la parte blanca del ojo y el interior del párpado\*. **(CUBRIR)**
1114. \*Al estornudar\*, cúbrete con la parte interna del codo. **(CUBRIR)**
1115. La bestia sale de noche, dijo Cuomo, refiriéndose a los escalofríos, dolores en el cuerpo y fiebre alta que \*lo\* visitaron el 1 de abril. **(VISITAR)**
1116. La situación puede empeorar si \*el virus\* deja el conducto bronquial y llega a los pulmones, donde causa una inflamación (neumonía). **(DEJAR)**
1117. \*Los límites extremos\* se han establecido entre 2 y 14 días después del contagio. **(ESTABLECER)**
1118. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió expertos \*a Wuhan\* para apoyar la investigación. **(ENVIAR)**
1119. Lo importante es que \*las personas\*, incluso si sufren de alergias, resfriados o gripe común, mantengan una distancia de un metro de los demás al toser o estornudar. **(MANTENER)**
1120. Una mujer china de 61 años vecina de Wuhan que el 8 de enero viajó en vuelo directo de Wuhan \*a Tailandia\* en compañía de cinco miembros de su familia en un grupo de dieciséis personas. **(VIAJAR)**
1121. Pon \*los guantes y las máscaras usados\* en un bote de basura con tapa en el cuarto de la persona enferma. **(PONER)**
1122. \*Los científicos\* pudieron advertir también que la anosmia -o pérdida del olfato- ya no se encuentra entre los 10 síntomas más comunes. **(ADVERTIR)**
1123. Un informe de la Organización Mundial de la Salud incluye el dolor de garganta \*entre los síntomas que pueden padecer los pacientes con coronavirus\*. **(INCLUIR)**
1124. Si atendemos \*al impacto socioeconómico y sociocultural de la pandemia /cierre de colegios, confinamiento, limitaciones de acceso a los servicios de salud\*, podremos observar que la pandemia de COVID-19 ha afectado de manera diferente a los géneros y posiblemente ha exacerbado las desigualdades de género-sexo. **(ATENDER)**
1125. \*Los mercados de Estados Unidos y de Alemania\* terminaron el año en máximos históricos, mientras que los de otros países como el Reino Unido y España cerraron 2020 con fuertes pérdidas. **(TERMINAR)**
1126. \*Aún no hemos llegado a ese punto, ya que todavía tenemos que probarlo en ensayos clínicos, pero es el objetivo final\*, afirmó Michael Snyder. **(AFIRMAR)**
1127. Pacientes recientemente hospitalizados que han sido ingresados en la unidad de cuidados intensivos (UCI) dentro de las 24 horas anteriores y que requieren \*ventilación mecánica invasiva, ventilación mecánica no invasiva (VNI), u oxígeno a alto flujo por cánula nasal de (HFNC) > 0,4 FiO<sub>2</sub> / 30 L / min de flujo de oxígeno. (BIIa)\***(REQUERIR)**
1128. Se está investigando el uso de donaciones de plasma sanguíneo de personas que ya se han recuperado de la COVID-19, \*una estrategia que\* se aplicó con éxito contra la gripe de 1918 y que se ha probado para el SARS, un primo anterior de COVID-19 y otras enfermedades como la fiebre hemorrágica argentina. **(APLICAR)**
1129. Tras el primer brote de COVID-19 en Wuhan en diciembre de 2019, donde las autoridades chinas confirmaron \*41 casos detectados entre el 8 de diciembre y el 2 de enero de 2020\*, la ciudad dejó de informar casos hasta el 19 de enero, cuando se confirmaron 17 casos más. **(CONFIRMAR)**

1130. Esto, de acuerdo con los hallazgos del estudio, suele ser un indicador de que los pacientes necesitan \*apoyo hospitalario\*. **(NECESITAR)**
1131. Las mascarillas normales buscan principalmente reducir la propagación de la enfermedad evitando contagiarse a otros, así que pueden dar \*una falsa sensación de seguridad\* a quienes las lleven. **(DAR)**
1132. Tras el primer brote de COVID-19 en Wuhan en diciembre de 2019, \*donde\* las autoridades chinas confirmaron 41 casos detectados entre el 8 de diciembre y el 2 de enero de 2020, la ciudad dejó de informar casos hasta el 19 de enero, cuando se confirmaron 17 casos más. **(CONFIRMAR)**
1133. Se han identificado casos en todos los continentes y, el 6 de marzo se confirmó el primer caso \*en Colombia\*. **(CONFIRMAR)**
1134. El 18 de marzo un artículo informa \*que el tratamiento con lopinavir/ritonavir da negativo en pruebas clínicas con 199 pacientes en China\*. **(INFORMAR)**
1135. \*Los NIH\* establecieron la Red de Prevención de COVID-19 (COVPN) para coordinar las redes de investigación existentes, a fin de llevar a cabo grandes estudios clínicos en decenas de miles de personas de manera eficiente. **(ESTABLECER)**
1136. Si tiene síntomas, pero obtiene un resultado negativo en la prueba de antígeno, su médico puede realizarle una prueba molecular para confirmar \*los resultados\*. **(CONFIRMAR)**
1137. El aislamiento se utiliza para separar a las personas que tienen el virus de la COVID-19 \*de las personas que no están enfermas\*. **(SEPARAR)**
1138. \*La persona que\* atiende al enfermo debe usar mascarilla y lavarse las manos tras cada contacto. **(ATENDER)**
1139. La mayoría de los casos de contagio se dan entre la población que tiene de 30 a 79 años, con una tasa de letalidad media del 2,3%, mientras que \*la mayor parte de los fallecidos\* supera la barrera de los 50 años, una edad a la que, según se deduce del mayor estudio elaborado hasta la fecha, aumenta el riesgo de mortalidad, con una tasa de letalidad del 1,3%. **(SUPERAR)**
1140. COVID-19 es una enfermedad altamente infecciosa que suele producir síntomas que \*una persona\* puede controlar en casa. **(CONTROLAR)**
1141. Algunas organizaciones internacionales, como la OMS, han publicado \*medidas preventivas\* para reducir la transmisión del virus. **(PUBLICAR)**
1142. El síndrome de liberación de citoquinas se produce cuando se activan \*grandes cantidades de leucocitos (neutrófilos, macrófagos y mastocitos)\* y liberan grandes cantidades de citoquinas proinflamatorias. **(ACTIVAR)**
1143. \*El paracetamol (Tylenol) y el ibuprofeno. (Advil, Motrin)\* ayudan a reducir la fiebre. **(AYUDAR)**
1144. La reducción de linfocitos B puede ocurrir temprano \*en la enfermedad\*, que puede afectar la producción de anticuerpos en el paciente. **(OCURRIR)**
1145. En tanto, quienes deseen hacerse el examen de forma voluntaria y están afiliados o afiliadas a Fonasa, podrán realizárselo \*en un hospital público\*. **(REALIZAR)**
1146. No obstante, actualmente parece ser que es más probable que una persona contagie cuando tiene síntomas, tales como fiebre o tos, pero se desconoce el periodo de tiempo durante el cual una persona con SARS puede transmitir la enfermedad \*a otras\*, es decir, podrían aparecer en tan solo 2 días. **(TRANSMITIR)**
1147. En un artículo publicado en la revista 'Science' dieron cuenta de algunos de los hallazgos que ya han hecho investigadores de todo el mundo y que ponen de manifiesto que \*la letalidad de la Covid-19\* se extiende prácticamente a todo el cuerpo, como se muestra en el gráfico. **(EXTENDER)**
1148. \*La prueba de antígeno\* también llamada prueba rápida da resultados en tan solo 15 minutos. **(DAR)**
1149. Las máscaras quirúrgicas (normales) son utilizadas desde hace tiempo por personas sanas \*en Hong Kong, Japón, Singapur y Malasia\*. **(UTILIZAR)**
1150. Ponte \*la mascarilla\*. **(PONER)**

1151. Causa una infección y, en ocasiones, al activarse \*el sistema inmunitario\*, este se descontrola y los pulmones pueden llenarse de líquido y células muertas. **(ACTIVAR)**
1152. La bestia sale de noche, dijo Cuomo, refiriéndose a \*los escalofríos, dolores en el cuerpo y fiebre alta que\* lo visitaron el 1 de abril. **(VISITAR)**
1153. Los síntomas pueden durar unos días, y \*quienes tienen el virus\* podrían sentirse mejor en más o menos una semana. **(SENTIR)**
1154. \*El coronavirus causante del covid-19\* se ha cobrado la vida de unas 286.000 personas –26.920 de ellas en España– y ha infectado a más de 4.177.000 en todo el mundo, según el último balance global ofrecido por la Universidad Johns Hopkins. **(INFECTAR)**
1155. Y los CDC y la OMS incluyen \*la diarrea\* entre los síntomas más comunes del COVID-19. **(INCLUIR)**
1156. Esto significa que \*la enfermedad\* se está extendiendo en varios países de todo el mundo, al mismo tiempo afectando a cientos de miles de personas en todos los continentes, causando un número de muertes que, a la fecha, sigue en ascenso. **(EXTENDER)**
1157. \*El virus\* entra por endocitosis. **(ENTRAR)**
1158. El nuevo coronavirus, que ahora se conoce con el nombre de 2019-nCoV o COVID-19, no se había detectado \*antes de que se notificara el brote en Wuhan (China) en diciembre de 2019\*. **(DETECTAR)**
1159. Esta afección dermatológica también fue uno de los primeros síntomas extraños que se identificaron \*tras el contagio con el SARS-CoV-2\*. **(IDENTIFICAR)**
1160. De acuerdo a la explicación de la OMS, el nuevo nombre se toma de las palabras corona, virus y disease (enfermedad en inglés), mientras que 19 representa el año en que surgió el brote se informó a la OMS \*el 31 de diciembre de 2019\*. **(INFORMAR)**
1161. \*Esto causa irritación y por ello empezamos a toser\*, señala Schaffner. **(SEÑALAR)**
1162. Cuando se la escribe con dígitos (COVID-19 en vez de COVID), el uso del guion es obligatorio en español (es incorrecto COVID19 –todo junto– o COVID 19 –con espacio), ya que lo posee en la denominación oficial, y, además, la lengua española tiene la costumbre de separar \*las palabras que se componen de letras y números\*. **(SEPARAR)**
1163. Después de que los investigadores prueban a fondo nuevos tratamientos o vacunas en el laboratorio para asegurarse de que puedan beneficiar a las personas, \*los tratamientos más prometedores\* pasan a los estudios clínicos. **(PASAR)**
1164. \*La Organización Mundial de la Salud\* recomienda por este motivo que se desinfecten las superficies, especialmente en el entorno sanitario. **(RECOMENDAR)**
1165. \*Una persona con fiebre\* sentirá que su espalda o pecho están calientes. **(SENTIR)**
1166. Las principales citoquinas implicadas en la patogénesis del CRS incluyen la interleuquina (IL)-6, la IL-10, el interferón (IFN), la proteína quimiotáctica de monocitos 1 (MCP1) y el factor estimulante de las colonias de granulocitos-macrófagos (GM-CSF); \*otras citoquinas como el factor de necrosis tumoral (TNF), IL-1, IL-2, IL-2-receptor- e IL-8\* también se han descrito durante el CRS. **(DESCRIBIR)**
1167. Sin embargo, con ómicron, la mialgia podría considerarse uno de los primeros indicios de que \*los pacientes\* se habían contagiado con la nueva versión del virus. **(CONTAGIAR)**
1168. La presencia de anticuerpos neutralizantes en la sangre se considera una prueba de inmunidad contra la infección, pero \*el nivel de anticuerpos neutralizantes\* disminuye con el tiempo, llegando a desaparecer tras tres meses en el 1% de los pacientes y tras seis meses en el 12%. **(DISMINUIR)**
1169. A pesar de que las autoridades del Servicio Nacional de Salud inglés (NHS) reconocen a la temperatura alta, tos persistente y a la pérdida o cambio repentino en el sentido de olfato o gusto. (anosmia y/o disgeusia) como principales síntomas del coronavirus, investigadores del King's College de Londres que dirigen \*el proyecto de vigilancia más ambicioso del mundo denominado COVID ZOE Symptom Study\* alertaron que la enfermedad pandémica está ahora actuando de manera diferente. **(DIRIGIR)**

1170. El 9 de febrero de 2021, una comisión internacional de la OMS enviada a China para investigar los orígenes de la pandemia \*in situ\* concluyó que el foco inicial de la pandemia fue el determinado desde un principio, el mercado mayorista de Wuhan. **(INVESTIGAR)**
1171. Más del 60 % de \*los 125 niños que\* se enfermaron gravemente o tuvieron una enfermedad crítica, tenían 5 años o menos. **(ENFERMAR)**
1172. Con mayor frecuencia, los síntomas aparecen \*alrededor de 5 días después de la exposición\*. **(APARECER)**
1173. Esta terapia fue autorizada \*por la FDA estadounidense\* en agosto de 2020 como tratamiento de emergencia contra la COVID-19. **(AUTORIZAR)**
1174. Las mujeres menores de 50 años deben ser conscientes del raro riesgo de que se formen coágulos de sangre \*después de la vacunación\*. **(FORMAR)**
1175. El 14 de febrero murió un turista chino \*en París\*, en la que se creyó se trataba de la primera muerte por COVID-19 fuera de Asia pero posteriormente se descubrió que un hombre fallecido en la ciudad española de Valencia el 13 de febrero también estaba infectado de coronavirus, que había contraído en Nepal. **(MORIR)**
1176. Para los pacientes hospitalizados con hipoxemia que requieren \*un suplemento de oxígeno convencional\*, el Panel recomienda usar una de las siguientes opciones: remdesivir (BIIa), dexametasona más remdesivir (BIII) o dexametasona sola (BI) (ver Manejo terapéutico de adultos con COVID-19). **(REQUERIR)**
1177. La Consejería de Sanidad considerará \*este resultado verdadero positivo\* y se recomendará autoaislamiento durante 7 días, evitar recibir visitas y extremar las medidas de prevención habituales. **(CONSIDERAR)**
1178. Mayo Clinic es una organización sin fines de lucro, y \*el dinero recaudado con la publicidad en Internet\* apoya nuestra misión. **(APOYAR)**
1179. Los estudios tienen como objetivo determinar si los inmunomoduladores pueden reducir \*la necesidad de ventiladores\* y acortar las estancias hospitalarias. **(REDUCIR)**
1180. Se presentan los casos acumulados que se han reportado \*en Puerto Rico\* desde el comienzo de la epidemia del SARS-CoV-2. **(REPORTAR)**
1181. \*En Estados Unidos\*, ya se está probando otra posible vacuna en humanos. **(PROBAR)**
1182. En tanto, quienes deseen hacerse el examen de forma voluntaria y están afiliados o afiliadas a Fonasa, podrán realizarse \*lo\* en un hospital público. **(REALIZAR)**
1183. Los científicos reconocen sentirse sorprendidos por el nivel y el tiempo que permanecen \*algunos de los cambios detectables en la sangre que\* la Covid-19 deja en el sistema inmunológico de los afectados, incluso en los casos de sintomatología leve. **(DEJAR)**
1184. El 2 de febrero de 2020, \*médicos de Tailandia\* afirmaron haber tratado a un paciente con éxito con una combinación de lopinavir / ritonavir y el medicamento contra la influenza oseltamivir. **(AFIRMAR)**
1185. \*El Centro Nacional de Desarrollo Biotecnológico de China\* afirmó el 17 de marzo que el antiviral Favipiravir, un inhibidor de la ARN polimerasa, mostró resultados positivos en un estudio de casos y controles con 80 pacientes en el Hospital Popular n.o 3 de Shenzhen, los que recibieron tratamiento con Favipiravir dieron negativo dentro de un periodo más reducido de tiempo en comparación con los del grupo de control, y recomienda que se incluya en el tratamiento. **(AFIRMAR)**
1186. Los investigadores han probado el nuevo algoritmo, desarrollado con software de código abierto, en 2.155 usuarios que durante ocho meses llevaron \*relojes inteligentes capaces de medir el estrés y las alteraciones en la frecuencia cardíaca\*, que si es más alta de lo normal puede ser señal clara de que existe algún problema. **(LLEVAR)**
1187. Se reportaron casos de reinfección en pacientes que \*ya\* habían pasado la enfermedad y se temía que el SARS-CoV-2 se comportase como algunos otros coronavirus, que son capaces de reinfectar después de aproximadamente un año. **(PASAR)**

1188. Tras analizar miles de casos de pacientes, los hospitales han determinado que existe un patrón de síntomas \*entre quienes se contagian de coronavirus\*. **(EXISTIR)**
1189. Los estudios realizados en Reino Unido, de seguimiento de síntomas desde los inicios de la pandemia, evidenciaron durante la ola de delta que los dolores de cabeza se habían convertido \*en el síntoma más común de la infección con la variante\*. **(CONVERTIR)**
1190. El virus ingresa a nuestro cuerpo y pasan \*varios días\* hasta que aparecen los primeros síntomas; cuáles son las señales para saber si estoy infectado. **(PASAR)**
1191. \*Schaffner\* compara la respuesta inflamatoria con un conflicto bélico. **(COMPARAR)**
1192. Pero ahora, la tendencia cambió y \*la tos\* ha sido superada por dolor de cabeza (66%), dolor de garganta (53%), secreción nasal (49%) y fiebre (42%) en menores de 40 años, mientras que los mayores de 40 mostraron resultados similares, precisó el profesor Spector. **(SUPERAR)**
1193. Bajo el punto de vista de uno de los virólogos españoles más reconocidos del mundo, la vacuna contra la gripe o los antivirales que tenemos no funcionan, así que son distintos, pero se transmiten de forma muy parecida y \*la enfermedad que causan\* también es similar, puede ir desde muy suave a muy severa. **(IR)**
1194. Los observadores han advertido que si bien la admiración de los trabajadores médicos de primera línea es generalizada y sincera, los medios estatales también deberían destacar \*la realidad de que muchos de esos trabajadores carecen de equipo de protección y que más de 3000 han sido infectados desde el brote y que la atención de los medios puede brindarles apoyo público para obtener algunos equipos muy necesarios\*. **(DESTACAR)**
1195. \*La Organización Mundial de la Salud (OMS) y la Organización Panamericana de la Salud (OPS)\* han reconocido el estado de pandemia, que supone el desarrollo de esta enfermedad. **(RECONOCER)**
1196. El SARS-CoV-2 puede propagarse de personas a animales, pero no se sabe con qué frecuencia sucede \*esto\*. **(SUCEDER)**
1197. Cúbrase \*la boca y la nariz\* con un pañuelo desechable o su manga (no sus manos) cuando tosa o estornude. **(CUBRIR)**
1198. El aislamiento ha provocado que \*muchos enfermos\* no puedan continuar con sus tratamientos, agravando sus problemas de salud. **(CONTINUAR)**
1199. El Centro Nacional de Desarrollo Biotecnológico de China afirmó \*el 17 de marzo\* que el antiviral Favipiravir, un inhibidor de la ARN polimerasa, mostró resultados positivos en un estudio de casos y controles con 80 pacientes en el Hospital Popular n.o 3 de Shenzhen, los que recibieron tratamiento con Favipiravir dieron negativo dentro de un periodo más reducido de tiempo en comparación con los del grupo de control, y recomienda que se incluya en el tratamiento. **(AFIRMAR)**
1200. Si tú o la persona que tiene COVID-19 presentan signos que advierten que se trata de una emergencia, se necesita \*atención médica inmediata\*. **(NECESITAR)**
1201. Si desarrolla cualquiera de los síntomas relacionado con el coronavirus, es importante que busque atención médica \*tan pronto como sea posible\*. **(BUSCAR)**
1202. Existe evidencia limitada que sugiere que el virus podría transmitirse uno o dos días antes de que se tengan síntomas, ya que \*la viremia\* alcanza un pico al final del período de incubación. **(ALCANZAR)**
1203. De acuerdo con los últimos datos del Ministerio de Sanidad, 28.993.301 personas en el país ya se encuentran inmunizadas, \*lo que\* representa el 61,1% de la población. **(REPRESENTAR)**
1204. El Sr. Chen agregó \*que si China no notificó a la OMS que era evidencia de un encubrimiento por parte de Beijing, o si China notificó a la OMS, entonces fue un incumplimiento del deber de la organización no transmitir información\*. **(AGREGAR)**
1205. Tal y como han informado los Institutos Nacionales de Salud (NIH, en inglés) a través de un comunicado, la fase 1 de la prueba clínica para evaluar una vacuna en investigación diseñada frente a la enfermedad del coronavirus ha empezado \*en el Instituto de Investigación Sanitaria Kaiser Permanente Washington, en Seattle\*, y se prevé que se alargue hasta seis semanas. **(EMPEZAR)**

1206. Las pruebas de laboratorio para el diagnóstico de la COVID-19 (enfermedad causada por el SARS-CoV-2) incluyen los métodos que detectan \*la presencia del propio virus\* y aquellos que detectan los anticuerpos producidos por el cuerpo humano en respuesta a la infección. **(DETECTAR)**
1207. A pesar de ello, los síntomas del contagio de la COVID-19 siguen manifestándose, \*cada vez\* de formas más variadas. **(MANIFESTAR)**
1208. Los institutos y centros de los NIH, en asociación con empresas privadas, continúan estudiando \*diferentes tipos de vacunas\*. **(ESTUDIAR)**
1209. La OMS recuerda que las mascarillas sólo son eficaces si se combinan con el lavado frecuente de manos con una solución hidroalcohólica con agua y jabón y únicamente recomienda su uso en personas sanas si estas atienden \*a alguien en quien se sospeche la infección\* o si se sufre tos y estornudos frecuentes. **(ATENDER)**
1210. Si la enfermedad no se trata \*a tiempo\*, estos síntomas pueden agravarse **(TRATAR)**
1211. Y los CDC y la OMS incluyen la diarrea \*entre los síntomas más comunes del COVID-19\*. **(INCLUIR)**
1212. Afirma que es imposible poner \*en un balanza\* los efectos adversos a medio y largo plazo. **(PONER)**
1213. Entre estos casos de covid-19 atribuidos a la variante ómicron, 34 (79%) ocurrieron en personas que completaron la serie primaria de una vacuna contra el covid-19 autorizada o aprobada por la FDA \*14 días antes del inicio de los síntomas o la recepción de una prueba positiva\*, incluidos 14 que habían recibido una dosis adicional o de refuerzo. **(COMPLETAR)**
1214. Desde el día 2 de marzo se incluye información sobre la evolución de la vacunación; \*los datos de vacunación\* no se actualizan los fines de semana y festivos. **(ACTUALIZAR)**
1215. \*Lo que\* compartimos en el ardor del momento depende de nosotros. **(COMPARTIR)**
1216. Se han identificado casos en todos los continentes y, \*el 6 de marzo\* se confirmó el primer caso en Colombia. **(CONFIRMAR)**
1217. El coronavirus se extendió por España \*desde febrero de 2020\* y a mediados de marzo obligó a imponer el estado de alarma para tratar de frenar los contagios y aliviar la saturación en los hospitales. **(EXTENDER)**
1218. Quizás debas considerar aislarte de la persona que está enferma y encontrar a \*otra persona que\* la cuide. **(CUIDAR)**
1219. Todo este proceso ocurrirá \*en el citoplasma de la célula\*. **(OCURRIR)**
1220. A finales de enero, \*la Comisión Nacional de Salud de China\* publicó una guía de manejo de las crisis psicológicas, en la que propugnaba la intervención de las personas afectadas, contactos cercanos, los encerrados en sus hogares, los familiares y amigos de los pacientes, personal sanitario y el público general que lo requiriera. **(PUBLICAR)**
1221. Suelen ser síntomas leves y comienzan gradualmente, por lo que en un principio es la fiebre y la tos lo primero que experimenta \*el infectado de COVID-19\*. **(EXPERIMENTAR)**
1222. Informó de que había visitado regularmente un mercado local de productos frescos \*en Wuhan\* antes de la aparición de la enfermedad el 5 de enero de 2020. **(VISITAR)**
1223. En lo que respecta al rumbo ambiental, las emisiones de gases de dióxido de nitrógeno cayeron \*a niveles mínimos históricos\* a 6 % a nivel mundial repartido entre las producciones cesadas en China y la Unión Europea. **(CAER)**
1224. Otros pacientes de Wuhan desarrollaron dificultad para respirar \*ese día\*. **(DESARROLLAR)**
1225. Con el frío llegan también \*los primeros resfriados y catarros\*. **(LLEGAR)**
1226. El virus que salió de Wuhan no es exactamente el mismo que entró \*en España\*. **(ENTRAR)**
1227. \*Albert\* sale de la UCI y pide a todos que se vacunen **(PEDIR)**
1228. El virus SARS-CoV-2 puede ingresar al cuerpo \*cuando otra persona inhala estas gotas\*. **(INGRESAR)**
1229. Todas las vacunas pueden causar algunos efectos secundarios, que son signos normales de que su cuerpo está produciendo \*anticuerpos\*. **(PRODUCIR)**
1230. En cambio, toma \*su temperatura\* al final de la tarde y al anochecer. **(TOMAR)**

1231. Exitosos libros como A Dancing Matrix (1994) de Robin Marantz Henig, The Coming Plague (1994) de Laurie Garrett y The Hot Zone de Richard Preston (1995) tocaron \*este tema\* y coincidieron en que mundialmente no existía la preparación suficiente para lidiar con algo así. **(TOCAR)**
1232. Esto, de acuerdo con los hallazgos del estudio, suele ser un indicador de que \*los pacientes\* necesitan apoyo hospitalario. **(NECESITAR)**
1233. La idea es perfeccionar las alertas para que puedan avisar de qué es lo que realmente está sucediendo y para ello planean agregar datos \*al algoritmo\* como los patrones de sueño, el recuento de pasos y la temperatura corporal. **(AGREGAR)**
1234. \*La inmensa mayoría\* coincide en los síntomas. **(COINCIDIR)**
1235. Se permiten \*visitas y acompañamiento\* a pacientes NO COVID en centros hospitalarios, bajo determinadas condiciones**(PERMITIR)**
1236. En cualquier caso, \*aquellos que porten este material\* deben aprender a usarlo y eliminarlo correctamente**(ELIMINAR)**
1237. De acuerdo a la explicación de la OMS, el nuevo nombre se toma de las palabras corona, virus y disease (enfermedad en inglés), mientras que 19 representa el año en que surgió \*el brote\* se informó a la OMS el 31 de diciembre de 2019. **(INFORMAR)**
1238. Lo que comenzó como un brote local en la provincia china de Hubei se convirtió \*en menos de tres meses\* en una pandemia global que afecta ya a casi todo el planeta. **(CONVERTIR)**
1239. Aunque las personas de avanzada edad son más vulnerables a la enfermedad, a mitad de marzo de 2020, \*la OMS\* recordó que los jóvenes también pueden tener complicaciones e incluso llegar a la muerte por el virus. **(RECORDAR)**
1240. El Nuevo Coronavirus COVID-19 es una cepa de la familia de coronavirus que no se había identificado \*previamente\* en humanos. **(IDENTIFICAR)**
1241. Más de 19.000 profesores y 100.000 alumnos permanecieron \*esta semana\* en cuarentena por covid. **(PERMANECER)**
1242. \*Además, la fiebre, tos seca y la pérdida del olfato y gusto, ya no son indicios tan comunes como al inicio de la pandemia, por ello esta lista de síntomas debe de cambiar\*, según manifiesta Tim Spector, líder del proyecto y profesor de epidemiología de la Universidad King's College de Londres. **(MANIFESTAR)**
1243. De acuerdo a un artículo publicado en The Conversation , los virus de la gripe y el SARS-CoV-2 que produce la COVID-19 no son lo mismo, por ello pueden coexistir, \*sus síntomas\* se parecen. **(PARECER)**
1244. Sin embargo, \*los adultos mayores y las personas de cualquier edad que tengan afecciones médicas previas\* deben llamar a su médico no bien aparecen los síntomas. **(LLAMAR)**
1245. El informe final se refiere \*al coronavirus\* como un nuevo patógeno que es altamente contagioso, puede propagarse rápidamente y debe considerarse capaz de causar enormes impactos en la salud, la economía y la sociedad. **(REFERIR)**
1246. A comienzos de 2020, el mundo dirigió la mirada hacia China, donde \*el SARS-CoV-2\* avanzaba por todo el país. **(AVANZAR)**
1247. La industria cinematográfica china había perdido \*US\$2 mil millones\* en marzo de 2020, después de haber cerrado todos sus cines durante el período del Año Nuevo Lunar que sustenta a la industria en toda Asia. **(PERDER)**
1248. \*La vacunación en España contra el coronavirus\* avanza a muy buen ritmo. **(AVANZAR)**
1249. Consulte \*los números de teléfono habilitados en su país\* para recibir atención médica y controlar la evolución de los síntomas. **(CONSULTAR)**
1250. El equipo encontró más casos relacionados al mercado y \*el 30 de diciembre\* las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió expertos a Wuhan para apoyar la investigación. **(COMUNICAR)**
1251. La llegada del nuevo coronavirus \*a América Latina\* ha generado dudas sobre la diferencia entre los síntomas de la enfermedad covid-19 y los de una gripe común. **(GENERAR)**

1252. Por ejemplo, se ha observado que las organizaciones de medios estatales People's Daily y Global Times, junto con el subdirector de información Zhao Lijian del Ministerio de Relaciones Exteriores, publican elogios efusivos sobre la respuesta de Beijing a la epidemia, como una amplia cobertura de la construcción acelerada de los nuevos hospitales en Wuhan (que Zhao afirmó se completó \*en 16 horas\*), la cuarentena de Wuhan con su población de 11 millones y la cuarentena sin precedentes de la provincia de Hubei. **(COMPLETAR)**
1253. Las principales citoquinas implicadas en la patogénesis del CRS incluyen la interleuquina (IL)-6, la IL-10, el interferón (IFN), la proteína quimiotáctica de monocitos 1 (MCP1) y el factor estimulante de las colonias de granulocitos-macrófagos (GM-CSF); otras citoquinas como el factor de necrosis tumoral (TNF), IL-1, IL-2, IL-2-receptor- e IL-8 también se han descrito \*durante el CRS\*. **(DESCRIBIR)**
1254. En un artículo publicado en la revista 'Science' dieron cuenta de algunos de los hallazgos que ya han hecho investigadores de todo el mundo y que ponen de manifiesto que la letalidad de la Covid-19 se extiende prácticamente \*a todo el cuerpo\*, como se muestra en el gráfico. **(EXTENDER)**
1255. Además, hay otros países como España y Bélgica que registran \*un número desigual entre hombres y mujeres\* durante toda la pandemia. **(REGISTRAR)**
1256. \*Las células B de memoria específicas para las proteínas de la nucleocápside y el pico del SARS-CoV-2\* duran al menos 6 meses después de la aparición de los síntomas. **(DURAR)**
1257. La enfermedad se propaga principalmente de persona a persona a través de las gotículas que salen despedidas \*de la nariz o la boca de una persona infectada\* al toser, estornudar o hablar. **(SALIR)**
1258. ¿O te estás cuidando a ti mismo \*en casa\*? **(CUIDAR)**
1259. \*El ácido acetilsalicílico. (aspirin)\* funciona bien para tratar la fiebre en adultos. **(FUNCIONAR)**
1260. La Consejería de Sanidad recuerda \*que hay que seguir manteniendo las medidas habituales de uso de mascarillas en interiores y en el exterior en lugares concurridos, ventilación de interiores, distancia de seguridad e higiene de manos\*. **(RECORDAR)**
1261. La Academia China de las Ciencias estimó \*en febrero de 2020\* para la COVID-19, un número reproductivo o  $R_0=4$ , aunque existe incertidumbre sobre muchos de los factores que se tomaron en consideración para calcular el  $R_0$ . **(ESTIMAR)**
1262. Algunos virus aéreos pueden acabar \*en el tracto digestivo\* y provocar problemas digestivos. **(ACABAR)**
1263. La mayoría de las personas tienen una enfermedad leve y logran recuperarse \*en casa\*. **(LOGRAR)**
1264. No obstante, estas medidas no son suficientes y urge ampliar los esfuerzos para mejorar \*la difusión de la disponibilidad de estos servicios\*. **(MEJORAR)**
1265. Quizás debas considerar aislarte de la persona que está enferma y encontrar a otra persona que \*la\* cuide. **(CUIDAR)**
1266. El 16 de marzo de 2020, \*una importante autoridad francesa y asesor del Gobierno francés sobre COVID-19, el profesor Didier Raoult del Instituto Universitario Hospitalario de Enfermedades Infecciosas (IHU-Méditerranée infection) en Marsella (Bouches-du-Rhône, Provenza-Alpes-Côte d'Azur)\*, anunció que un ensayo con 24 pacientes del sureste de Francia había demostrado que la cloroquina es un tratamiento efectivo para COVID-19. **(ANUNCIAR)**
1267. A comienzos de 2020, \*el mundo\* dirigió la mirada hacia China, donde el SARS-CoV-2 avanzaba por todo el país. **(DIRIGIR)**
1268. En un documento interno obtenido por el New York Times, los CDC indican que \*la mayor propagación de casos con pronósticos severos\* está ocurriendo en lugares con tasas bajas de vacunación. **(OCURRIR)**
1269. \*Los desechos\* deben ponerse en una bolsa plástica, cerrarla y botarla. **(PONER)**
1270. Asimismo, se ha confirmado que otros animales –como perros, gatos, tigres, leones y murciélagos– pueden contraer \*el COVID-19\* al igual que los seres humanos. **(CONTRAER)**
1271. A finales de marzo, Estados Unidos se convirtió \*en el país más afectado del mundo\*. **(CONVERTIR)**

1272. Se presentan \*los casos acumulados que\* se han reportado en Puerto Rico desde el comienzo de la epidemia del SARS-CoV-2. **(REPORTAR)**
1273. Signos que advierten \*que se trata de una emergencia\***(ADVERTIR)**
1274. \*La industria cinematográfica china\* había perdido US\$2 mil millones en marzo de 2020, después de haber cerrado todos sus cines durante el período del Año Nuevo Lunar que sustenta a la industria en toda Asia. **(PERDER)**
1275. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió expertos a Wuhan para apoyar \*la investigación\*. **(APOYAR)**
1276. \*Los síntomas\* comienzan de forma leve. **(COMENZAR)**
1277. Al igual que el sistema sanitario está saturado, los refugios y las líneas de atención para las víctimas de violencia doméstica han alcanzado \*su límite de capacidad\*, agravándose además por el hecho de haber tenido que adaptar estos lugares para sosegar el colapso hospitalario. **(ALCANZAR)**
1278. \*Los padres y cuidadores de niños que presentan señales de una infección por coronavirus\* deben comunicarse con su médico y mantener al niño en casa. **(COMUNICAR)**
1279. \*Las personas que\* contraen el coronavirus también pueden tener diarrea, como suele ocurrir en los niños que cogen la gripe. **(CONTRAER)**
1280. Varias empresas farmacéuticas, como Vir Biotechnology, han evaluado \*la efectividad contra el virus SARS-Cov-2 de anticuerpos monoclonales previamente identificados contra otros virus\*. **(EVALUAR)**
1281. Si solo estás sufriendo \*dolor de cabeza\* sin ningún otro síntoma (fiebre, tos...) probablemente sea migraña o dolor tensional. **(SUFRIR)**
1282. Debe permanecer en casa y evitar el contacto con personas hasta que su proveedor le indique que es seguro terminar \*el aislamiento\* en casa. **(TERMINAR)**
1283. \*Hacerlo frecuentemente\* reduce hasta en 50% el riesgo de contraer coronavirus. **(REDUCIR)**
1284. Asegurarse de que \*las vacunas y los tratamientos contra el COVID-19\* funcionen para la mayor cantidad de personas posible. **(FUNCIONAR)**
1285. Debido a la complejidad y coste de las pruebas PCR, en los primeros meses de la pandemia \*casi ningún país\* dispuso de datos confiables sobre la prevalencia del virus en su población. **(DISPONER)**
1286. Durante los primeros meses, cuando \*la epidemia\* se limitaba casi exclusivamente a China, se produjo escasez generalizada de productos farmacéuticos, electrónicos y otros productos manufacturados debido a la paralización de numerosas fábricas en China. **(LIMITAR)**
1287. Así mismo, tanto como el Festival de la Canción de Eurovisión 2020, por primera vez en su historia, y \*otros eventos deportivos mundiales como la Eurocopa 2020, la Copa América 2020, la Temporada 2020 de Fórmula 1 y los Juegos Olímpicos de Tokio 2020\*, han sido cancelados o aplazados hasta este año 2021, entre varios otros eventos mundiales**(CANCELAR)**
1288. Mantén \*una rutina diaria\*, que incluya bañarte, o ducharte, y vestirte. **(MANTENER)**
1289. Hay otros signos del coronavirus que no presentan \*los enfermos de gripe\*, como la pérdida del gusto o el olfato, las erupciones cutáneas, las diarreas, los cambios de color en los dedos de las manos o los pies. **(PRESENTAR)**
1290. Existe \*mucha información errónea\* en línea sobre los tratamientos para el COVID-19. **(EXISTIR)**
1291. Quienes ya se estaban recuperando pueden ver que sus fiebres desaparecen \*después de 12 días\*. **(DESAPARECER)**
1292. \*Cada virus\* puede crear entre 10.000 y 100.000 réplicas. **(CREAR)**
1293. Pudimos determinar \*que los pacientes que no habían recibido Plaquenil (el medicamento que contiene hidroxyclorequina) seguían siendo contagiosos después de seis días, pero de los que habían recibido Plaquenil, después de seis días, solo el 25% seguía siendo contagioso\*. **(DETERMINAR)**
1294. Ante la incertidumbre generada, la OMS ha desmentido que se propague a grandes distancias a través del aire y ha asegurado \*que la principal vía de contagio son las gotículas que se generan

- cuando una persona infectada tose o estornuda, o a través de saliva o de secreciones de la nariz.  
(ASEGURAR)
1295. \*El 7 de enero de 2020\* los científicos chinos habían aislado el virus causante de la enfermedad, y realizaron la secuenciación del genoma. (REALIZAR)
1296. Desde el 24 de marzo de 2020 esta prueba pasó a ser gratuita para \*quienes\* pertenecen a cualquier tramo. (A, B, C o D). (PERTENECER)
1297. Los nuevos datos apuntan \*a que no es el coronavirus el que ha cambiado radicalmente, sino la humanidad, gracias a las vacunas y a las defensas adquiridas en infecciones previas\*(APUNTAR)
1298. En China se observó también \*que la mortalidad aumentaba con la edad y es más alta entre los hombres que en las mujeres (4,7 % frente al 2,8 %)\*. (OBSERVAR)
1299. \*La Academia China de las Ciencias\* estimó en febrero de 2020 para la COVID-19, un número reproductivo o  $R_0=4$ , aunque existe incertidumbre sobre muchos de los factores que se tomaron en consideración para calcular el  $R_0$ . (ESTIMAR)
1300. Lo importante es que las personas, incluso si sufren de alergias, resfriados o gripe común, mantengan \*una distancia de un metro de los demás\* al toser o estornudar. (MANTENER)
1301. Así, \*los firmantes\* piden que se amplíe el cuadro clínico porque si no, el país podría pasar por alto o retrasar la identificación de muchos casos de Covid. (PEDIR)
1302. ¿Se siguen investigando \*las vacunas\*? (INVESTIGAR)
1303. Los CDC ofrecen \*una guía\* para el tratamiento de COVID-19 en el hogar. (OFRECER)
1304. El nuevo coronavirus, que ahora se conoce con el nombre de 2019-nCoV o COVID-19, no se había detectado antes de que se notificara \*el brote en Wuhan (China)\* en diciembre de 2019. (NOTIFICAR)
1305. En cambio, los respiradores certificados también dan una gran protección \*a su portador\*. (DAR)
1306. A finales de enero, \*el Ministerio de Salud de Rusia\* identificó tres medicamentos para adultos que podrían ayudar a tratar la enfermedad. (IDENTIFICAR)
1307. La de 'tipo L' presentaba especial virulencia y fue más común al comienzo del brote, pero su frecuencia ha disminuido \*después de principios de enero\*, a causa, según los científicos, de las estrictas medidas que China puso en marcha. (DISMINUIR)
1308. Si bien \*las medidas de contención\* previenen la propagación del virus, las mujeres y niñas se ven obligadas a convivir con personas que son violentas y quedan aisladas de los recursos que pueden ayudarlas. (PREVENIR)
1309. Así pues, las mujeres se enfrentan a \*una carga triple de trabajo que\* puede agravar las desigualdades preexistentes. (AGRAVAR)
1310. Pese a que \*muchas personas\* han comparado este virus con el de la gripe, la OMS cifra en un 3,4% la proporción de casos que terminan con la muerte del paciente. (COMPARAR)
1311. Especialistas apuntan a que esta tendencia se debe a que \*nuestra sociedad\* sigue asociando a las mujeres a las labores de cuidados, aunque sea de forma inconsciente. (ASOCIAR)
1312. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al \*CCDC, que\* envió expertos a Wuhan para apoyar la investigación. (ENVIAR)
1313. Los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) ofrecen \*más información sobre la autoevaluación\*, por ejemplo, cómo hacerse un hisopado. (OFRECER)
1314. Y aunque esas consecuencias no resulten muy dañinas, lo cierto es que pueden ser muy duraderas, ya que los investigadores hablan, incluso, de la existencia de secuelas incluso diez semanas después de que el virus SARS-CoV-2 haya desaparecido \*del organismo del paciente\*. (DESAPARECER)
1315. Se administraron \*600 mg de hidroxiclороquina (marca Plaquenil)\* a estos pacientes todos los días durante 10 días. (ADMINISTRAR)
1316. \*Los investigadores\* han descubierto que el virus puede permanecer en el aire durante, al menos, 30 minutos y viajar hasta 4,5 metros. (DESCUBRIR)

1317. Sin embargo, la ausencia de anticuerpos en la sangre no significa que \*el sistema inmunitario\* no pueda producir nuevos anticuerpos rápidamente en caso de reexposición al SARS-CoV-2. **(PRODUCIR)**
1318. Al igual que el sistema sanitario está saturado, \*los refugios y las líneas de atención para las víctimas de violencia doméstica\* han alcanzado su límite de capacidad, agravándose además por el hecho de haber tenido que adaptar estos lugares para sosegar el colapso hospitalario. **(ALCANZAR)**
1319. Durante más de dos meses, los españoles sufrieron \*uno de los confinamientos más estrictos\*, lo que no evitó que los servicios sanitarios se vieran superados en muchos lugares. **(SUFRIR)**
1320. Según la Superintendencia de Salud, las Isapres al ser parte de la seguridad social y de la protección del derecho a la salud, no podrán, en ningún caso, excluir \*la cobertura a la enfermedad\* aunque haya sido declarada pandemia. **(EXCLUIR)**
1321. Tome \*la cantidad recomendada\* para reducir la fiebre. **(TOMAR)**
1322. Los casos crecieron rápidamente y el 13 de marzo de 2020 la Organización Mundial de la Salud declaró a Europa el nuevo epicentro del virus después de que la situación mejorara \*en China\*. **(MEJORAR)**
1323. Especialistas apuntan a que esta tendencia se debe a que nuestra sociedad sigue asociando \*a las mujeres\* a las labores de cuidados, aunque sea de forma inconsciente. **(ASOCIAR)**
1324. \*La partícula viral\* viajará, a través del sistema de transporte de vesículas celular en el que interviene el aparato de Golgi, hasta la superficie. **(VIAJAR)**
1325. \*El Ministerio de Defensa de China\* asegura haber desarrollado con éxito una vacuna contra el coronavirus y ha autorizado las pruebas en humanos. **(ASEGURAR)**
1326. Los datos de este estudio reflejan que apenas presentan esta característica el 13,9% de los afectados, pero sí puede ser un indicador que advierta de \*la presencia del virus en nuestro organismo\*. **(ADVERTIR)**
1327. \*El Sindicat d'Habitatge de Nou Barris\* cree que se están acumulando los desalojos en previsión de que acabe la moratoria. **(CREER)**
1328. Se conoce que cualquier persona puede infectarse, independientemente de su edad, pero hasta el momento se han registrado relativamente \*pocos casos de COVID-19 en niños\*. **(REGISTRAR)**
1329. Los médicos pueden determinar qué medidas tomar para cuidar al paciente, y pueden evitar que \*la condición del paciente\* empeore, resalta uno de los líderes de la investigación, Peter Kuhn. **(EMPEORAR)**
1330. Los científicos evalúan \*la seguridad de los regímenes potenciados mixtos\* y controlan los niveles de anticuerpos resultantes. **(EVALUAR)**
1331. \*Todo este proceso\* ocurrirá en el citoplasma de la célula. **(OCURRIR)**
1332. Asegúrate de que en casa todos conocen \*cómo hacer un correcto lavado de manos\*. **(CONOCER)**
1333. Así \*lo\* hace cada comunidad. **(HACER)**
1334. Además, \*los factores inflamatorios asociados con las enfermedades que contienen principalmente IL-6\* aumentaron significativamente, lo que también contribuyó al agravamiento de la enfermedad alrededor de 2 a 10 días después del inicio. **(AUMENTAR)**
1335. Tanto en este último caso como en el de aquellas que trabajan en la economía sumergida, la crisis sanitaria causada por el coronavirus ha provocado que muchas pierdan sus puestos de trabajos, tal y como ha sucedido \*en la mayoría de los países\*. **(SUCEDER)**
1336. En un documento interno obtenido por el New York Times, los CDC indican que la mayor propagación de casos con pronósticos severos está ocurriendo \*en lugares con tasas bajas de vacunación\*. **(OCURRIR)**
1337. \*La novedad\* surge del rastreo que hace la aplicación COVID ZOE Symptom Study, el mayor monitoreo comunitario del mundo. **(SURGIR)**
1338. Se obtuvieron \*muestras de estos pacientes\* para realizar análisis de laboratorio. **(OBTENER)**
1339. \*La FDA\* ha aprobado o autorizado varios tratamientos para enfermedades leves, moderadas y graves causadas por el COVID-19. **(APROBAR)**

1340. Los mercados de Estados Unidos y de Alemania terminaron \*el año\* en máximos históricos, mientras que los de otros países como el Reino Unido y España cerraron 2020 con fuertes pérdidas. **(TERMINAR)**
1341. El virus comienza \*un segundo ataque\*, causando que la condición del paciente empeorar alrededor de 7 a 14 días después del inicio. **(COMENZAR)**
1342. Tenían una ventaja porque ya sabían mucho \*sobre cómo funcionan las vacunas de ARNm y sobre cómo fabricarlas\*. **(SABER)**
1343. La mayoría de los proveedores de atención médica ofrecen \*pruebas de COVID-19\*, y debe contactarlos primero a ellos si tiene síntomas. **(OFRECER)**
1344. El gobierno federal está proporcionando \*vacunas contra el COVID-19 autorizadas por la FDA\* de forma gratuita a todas las personas que viven en los Estados Unidos, independientemente de su estado migratorio o de seguro médico. **(PROPORCIONAR)**
1345. Su proveedor puede tomar muestras \*con hisopos\* de la parte posterior de su nariz y garganta. **(TOMAR)**
1346. Todas las vacunas pueden causar algunos efectos secundarios, que son signos normales de que \*su cuerpo\* está produciendo anticuerpos. **(PRODUCIR)**
1347. \*Lo que comenzó como un brote local en la provincia china de Hubei\* se convirtió en menos de tres meses en una pandemia global que afecta ya a casi todo el planeta. **(CONVERTIR)**
1348. \*El coronavirus\* se extendió por España desde febrero de 2020 y a mediados de marzo obligó a imponer el estado de alarma para tratar de frenar los contagios y aliviar la saturación en los hospitales. **(EXTENDER)**
1349. Los estudios realizados en Reino Unido, de seguimiento de síntomas desde los inicios de la pandemia, evidenciaron \*durante la ola de delta\* que los dolores de cabeza se habían convertido en el síntoma más común de la infección con la variante. **(EVIDENCIAR)**
1350. Por otro lado, la pandemia ha afectado \*a la salud mental, social y sexual de las personas LGTBI\* por varios motivos. **(AFECTAR)**
1351. Sin embargo, los presuntos casos de COVID-19 en \*las dos provincias chinas (Liaoning y Jilin) que\* limitan con Corea del Norte han sido bajos. **(LIMITAR)**
1352. \*Algunos tratamientos propuestos\* no han demostrado ser efectivos y podrían ser peligrosos. **(DEMOSTRAR)**
1353. Debe permanecer en casa y evitar el contacto con personas hasta que su proveedor \*le\* indique que es seguro terminar el aislamiento en casa. **(INDICAR)**
1354. Actualmente, estamos viviendo \*una tercera ola con datos preocupantes\* en todo el mundo y un diagnóstico prematuro puede evitar contagios de las personas más vulnerables al virus (personas mayores, afectados por patologías previas...) si se guarda de manera estricta el aislamiento de dos semanas posterior a la detección. **(VIVIR)**
1355. De todos modos, la duración y la gravedad de los síntomas depende \*de si la persona está vacunada o no\*. **(DEPENDER)**
1356. \*Corea del Norte sigue siendo el único país en el que\* aún se desconoce algún caso confirmado de COVID-19, aunque fuentes del gobierno de Corea del Sur sospechan que la enfermedad ha afectado al país. **(DESCONOCER)**
1357. \*Casi la totalidad de instituciones educativas de todos los niveles\* cancelaron sus actividades presenciales, en cumplimiento de las disposiciones adoptadas por las autoridades de los distintos países, tendientes a disminuir la propagación de la pandemia y sus consecuencias fatales. **(CANCELAR)**
1358. Pese a que en la gran mayoría de los casos \*estos síntomas\* coinciden con otras patologías que no son la COVID-19, hoy por hoy sin hacer un test del coronavirus no hay seguridad plena. **(COINCIDIR)**
1359. Para que la alarma saltara, \*la frecuencia cardíaca\* debía elevarse durante un espacio prolongado de tiempo, por lo que si alguno salía a correr no saltaba. **(ELEVAR)**

1360. Llama al 911 o al número local para emergencias si no puedes despertar a la persona enferma o si notas signos que indiquen \*una emergencia\*. **(INDICAR)**
1361. \*Los datos\* han demostrado que la sintomatología del coronavirus puede ser leve, similar al de un catarro suave. **(DEMOSTRAR)**
1362. \*La persona encargada de los cuidados u otra persona\* debería usar una mascarilla y guantes desechables antes de limpiar. **(USAR)**
1363. A finales de enero, la Comisión Nacional de Salud de China publicó \*una guía de manejo de las crisis psicológicas\*, en la que propugnaba la intervención de las personas afectadas, contactos cercanos, los encerrados en sus hogares, los familiares y amigos de los pacientes, personal sanitario y el público general que lo requiriera. **(PUBLICAR)**
1364. \*Hasta el momento\*, las pruebas que se han realizado descartan la presencia de SARS-CoV-2 en el líquido amniótico, en la sangre de cordón umbilical y en la leche materna. **(DESCARTAR)**
1365. También se excluyen \*casos de las repúblicas populares no reconocidas de Donetsk y Lugansk\*. **(EXCLUIR)**
1366. Además de la guía mencionada anteriormente sobre el lavado de manos y la higiene respiratoria, los organismos de salud pública aconsejan \*que las personas enfermas que sospechan que pueden tener COVID-19 deben restringir las actividades fuera del hogar, excepto para obtener atención médica\*. **(ACONSEJAR)**
1367. Cuando \*el virus\* entra en nuestro cuerpo -ya sea por los ojos, la boca o la nariz- se sujeta a las células de la mucosa del fondo de la nariz y la garganta, dice el experto. **(ENTRAR)**
1368. Estos estudios se centran en personas a las que se les diagnostica \*COVID-19\* y están hospitalizadas, así como en aquellas que se están recuperando de COVID-19. **(DIAGNOSTICAR)**
1369. A veces, las personas que están completamente vacunadas contraen \*una infección irruptiva\*, lo que significa que dan positivo en la prueba del SARS-CoV-2 o se enferman con COVID-19. **(CONTRAER)**
1370. Comparte esta información con tu familia, tus amigos y tus colegas para conseguir que todo el mundo conozca \*todos los datos relativos al COVID-19\* y puedan proteger su salud. **(CONOCER)**
1371. \*Los investigadores\* creen que la variante Delta proveniente de la India podría estar detrás del cambio en el cuadro clínico. **(CREER)**
1372. El Instituto Feinstein de Northwell Health anunció en marzo un estudio sobre \*un anticuerpo humano que\* puede prevenir la actividad de IL-6. **(PREVENIR)**
1373. Es lo que puede ocurrir cuando \*la respuesta inflamatoria es tan vigorosa que\* se suma al problema de la neumonía. **(SUMAR)**
1374. El COVID-19 ha demostrado, desde las primeras olas de la pandemia, \*que las enfermedades físicas también pueden causar afecciones mentales\*. **(DEMOSTRAR)**
1375. La infección se produce cuando una persona enferma tose o estornuda y expulsa partículas del virus que entran \*en contacto\* con otras personas. **(ENTRAR)**
1376. Conforme a las indicaciones de los Centros para el Control y Prevención de Enfermedades (CDC, por sus siglas en inglés), los pacientes podrían experimentar \*náuseas o vómitos\*. **(EXPERIMENTAR)**
1377. Unos investigadores han creado \*un algoritmo que\* permite a los relojes inteligentes tener la capacidad de detectar si tiene síntomas de la Covid-19. **(PERMITIR)**
1378. Sanidad aprueba \*una cuarta dosis de la vacuna\* para personas inmunodeprimidas cinco meses después de la tercera. **(APROBAR)**
1379. Los virus de la gripe y la COVID-19 provocan \*enfermedades respiratorias\*, por lo que algunos de sus síntomas más característicos son similares: fiebre, tos, fatiga, dificultad para respirar, dolor de garganta, mucosidad nasal, dolor muscular y dolor de cabeza. **(PROVOCAR)**
1380. \*Desde principios de mayo\*, hemos estado observando los principales síntomas de los usuarios de la aplicación, y no son los mismos que antes, explica el experto. **(OBSERVAR)**
1381. Lo que compartimos en el ardor del momento depende \*de nosotros\*. **(DEPENDER)**

1382. Otros pacientes podrían haber sufrido esta afección, pero no se incluye entre las afecciones que \*la Organización Mundial de la Salud (OMS)\* asocia al COVID-19. **(ASOCIAR)**
1383. Controla atentamente \*los síntomas, tuyos o de tu ser querido\*, para notar si están empeorando. **(CONTROLAR)**
1384. El Ministerio de Educación ha publicado \*un balance con la incidencia reportada por las CC AA el 12 de enero\*. **(PUBLICAR)**
1385. La llegada del nuevo coronavirus a América Latina ha generado \*dudas sobre la diferencia entre los síntomas de la enfermedad covid-19 y los de una gripe común\*. **(GENERAR)**
1386. Presentaron \*linfopenia\* 26 de 41 pacientes (63 %). **(PRESENTAR)**
1387. Los NIH están patrocinando estudios para probar la seguridad y eficacia de los inmunomoduladores, que minimizan los efectos de una respuesta hiperactiva del sistema inmunológico que experimentan \*algunos pacientes con COVID-19\*. **(EXPERIMENTAR)**
1388. La arquitecta Carmen Romero propuso llenar el Oso y el Madroño de estrellas, en recuerdo de los que \*nos\* dejaron a causa de la covid**(DEJAR)**
1389. La infección se produce cuando \*una persona enferma\* tose o estornuda y expulsa partículas del virus que entran en contacto con otras personas. **(TOSER)**
1390. \*La OMS\* actualiza los indicios para descubrir si estás contagiado; consulta en el gráfico todas las partes del cuerpo a las que puede afectar el virus**(ACTUALIZAR)**
1391. A lo largo de todo el 2020 y en los meses transcurridos de 2021 \*la información disponible acerca del coronavirus y su forma de transmisión\* fue variando. **(VARIAR)**
1392. \*La reducción de linfocitos B\* puede ocurrir temprano en la enfermedad, que puede afectar la producción de anticuerpos en el paciente. **(OCURRIR)**
1393. Sin embargo, la ausencia de anticuerpos en la sangre no significa \*que el sistema inmunitario no pueda producir nuevos anticuerpos rápidamente en caso de reexposición al SARS-CoV-2\*. **(SIGNIFICAR)**
1394. Estas pequeñas estructuras son las que normalmente se llenan de aire, y a través de sus paredes se produce el intercambio gaseoso por el cual \*el oxígeno\* llega a la sangre, y de allí al resto del cuerpo. **(LLEGAR)**
1395. Los voluntarios adultos que han sido completamente vacunados contra el COVID-19 reciben \*dosis de refuerzo de diferentes vacunas contra el COVID-19\*. **(RECIBIR)**
1396. El Sindicat d'Habitatge de Nou Barris cree \*que se están acumulando los desalojos en previsión de que acabe la moratoria\*. **(CREER)**
1397. Además, \*este tipo de medidas\* han agravado la situación de soledad en miembros de la comunidad, particularmente en personas mayores LGTBI, ya que frecuentemente viven solas y no tienen relaciones positivas con sus familias. **(AGRAVAR)**
1398. Además, también alertó de un periodo de estrés después de que muchos de los participantes recibieran \*la vacuna del coronavirus\*, que podría deberse al aumento en la respuesta inmune generada por el pinchazo. **(RECIBIR)**
1399. \*Todos quienes\* viven con el paciente deben ser monitoreados por la Autoridad Sanitaria. **(VIVIR)**
1400. No obstante, este tiempo de permanencia depende \*de factores como la temperatura y el tipo de superficie\*. **(DEPENDER)**
1401. Sin embargo, los presuntos casos de COVID-19 en las dos provincias chinas (Liaoning y Jilin) que limitan \*con Corea del Norte\* han sido bajos. **(LIMITAR)**
1402. Si la situación del paciente sigue agravándose, este es el momento en el que es probable que ingresen \*a la unidad de cuidados intensivos (UCI)\*. **(INGRESAR)**
1403. La pandemia de COVID-19 en América es una epidemia continental que inició con la detección del primer caso de esta enfermedad el 21 de enero de 2020 en los Estados Unidos, un hombre de aproximadamente 30 años, originario del estado de Washington, que había viajado \*recientemente\* a China. **(VIAJAR)**

1404. Además, también alertó de un periodo de estrés después de que \*muchos de los participantes\* recibieran la vacuna del coronavirus, que podría deberse al aumento en la respuesta inmune generada por el pinchazo. **(RECIBIR)**
1405. El Ministerio de Defensa de China asegura \*haber desarrollado con éxito una vacuna contra el coronavirus\* y ha autorizado las pruebas en humanos. **(ASEGURAR)**
1406. El español cree \*que toda la polémica alrededor de Novak Djokovic no es buena para el tenis\*. **(CREER)**
1407. \*La situación de vulnerabilidad estructural del colectivo LGTBI\* también ha incrementado a raíz de la pandemia. **(INCREMENTAR)**
1408. El equipo encontró \*más casos relacionados al mercado\* y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos al CCDC, que envió expertos a Wuhan para apoyar la investigación. **(ENCONTRAR)**
1409. El equipo encontró más casos relacionados al mercado y el 30 de diciembre las autoridades de salud de Wuhan comunicaron los casos \*al CCDC\*, que envió expertos a Wuhan para apoyar la investigación. **(COMUNICAR)**
1410. \*Los pacientes que\* pierden el olfato no siempre pierden el gusto. **(PERDER)**
1411. A esto le sumamos que debido al teletrabajo y demanda de trabajos a tiempo parcial ha aumentado y está más solicitado por \*mujeres que\* tienen que cuidar de sus hijos y hacer las tareas domésticas al mismo tiempo. **(CUIDAR)**
1412. \*Las vacunas\* protegen a las personas de infectarse o de enfermarse gravemente, y las mascarillas y el distanciamiento social ayudan a evitar que el virus se propague **(PROTEGER)**
1413. \*Los expertos\* no ven necesaria la cuarta dosis para todos **(VER)**
1414. La Academia China de las Ciencias estimó en febrero de 2020 para la COVID-19, \*un número reproductivo o  $R_0=4$ \*, aunque existe incertidumbre sobre muchos de los factores que se tomaron en consideración para calcular el  $R_0$ . **(ESTIMAR)**
1415. Respecto a la aparición de los síntomas tras el contagio, \*la OMS\* destaca que la media es de 5 o 6 días, aunque dependiendo de la cepa pueden aparecer hasta los 14 días. **(DESTACAR)**
1416. Personas que llevan \*máscaras quirúrgicas\* en Guangzhou. **(LLEVAR)**
1417. \*Si no tienes seguro, puedes comenzar con el departamento de salud estatal o los centros de salud comunitarios locales, que se conocen oficialmente como centros de salud calificados federalmente\*, aconsejó Harris, y agregó que algunos estados tienen un número de teléfono directo 1-800 para llamar. **(ACONSEJAR)**
1418. Cámbiate la mascarilla \*todos los días\*. **(CAMBIAR)**
1419. \*El proyecto Zoe\* también demostró que las personas que ya han sido vacunadas y dieron positivo estornudaban con más frecuencia. **(DEMOSTRAR)**
1420. \*Las complicaciones\* incluían el síndrome de dificultad respiratoria aguda (resultado positivo para RT-PCR en tiempo real en la muestra de plasma), lesión cardíaca aguda, e infección secundaria. **(INCLUIR)**
1421. \*El Ministerio de Educación\* ha publicado un balance con la incidencia reportada por las CC AA el 12 de enero. **(PUBLICAR)**
1422. \*Un estudio con 2143 niños\* sugiere que este grupo de la población puede ser un factor crítico en la rápida propagación de la enfermedad. **(SUGERIR)**
1423. El proyecto Zoe ha demostrado que \*la forma en la que el coronavirus se manifiesta en los contagiados\* ha variado en gran medida con la evolución de la pandemia. **(VARIAR)**
1424. Los ciudadanos podrán retirarlos con su Tarjeta Sanitaria o, si pertenecen \*a alguna mutua\* y no cuentan con tarjeta sanitaria pública, presentando su DNI o NIE. **(PERTENECER)**
1425. Investigadores especializados en enfermedades infecciosas de la Universidad de Texas en Austin (Estados Unidos) y otras instituciones de Hong Kong, China y Francia han llegado a la conclusión de que es muy probable que el coronavirus SARS-CoV-2 se propagara más allá de Wuhan y otras ciudades antes de que las autoridades chinas pudieran activar \*las medidas de cuarentena\*. **(ACTIVAR)**

1426. Llame \*antes de recibir atención médica\*. **(LLAMAR)**
1427. \*Los CDC\* han actualizado las recomendaciones para el público sobre el aislamiento y la cuarentena y están revisando el sitio web de los CDC para incluir estos cambios. **(REVISAR)**
1428. A finales de mayo, América pasa a ser el epicentro mundial de la pandemia del coronavirus, pero la llegada de las vacunas en 2021 y su administración en los países más desarrollados ha hecho aparecer \*una mayor transmisión relativa de la COVID-19\* en Asia. **(APARECER)**
1429. Si estás cuidando a alguien con COVID-19 y no estás completamente vacunado, los CDC recomiendan \*que hagas cuarentena\* durante 14 días después del último contacto con la persona enferma y que estés atento a cualquier síntoma de la COVID-19. **(RECOMENDAR)**
1430. \*El 18 de diciembre de 2020\*, la FDA autorizó el uso de emergencia de esta vacuna contra el COVID-19 financiada por los NIH en los Estados Unidos para personas mayores de 18 años. **(AUTORIZAR)**
1431. \*Personas que\* llevan máscaras quirúrgicas en Guangzhou. **(LLEVAR)**
1432. \*Investigadores de la Universidad Noruega de Ciencia y Tecnología (NTNU)\* han creado una base de datos con 120 agentes antivirales de amplio espectro seguros para las personas e identificaron 31 candidatos a fármacos para el tratamiento de SARS-CoV-2. **(CREAR)**
1433. La Organización apoya \*a los países\* en la vigilancia, la prevención y la preparación para el control de enfermedades pandémicas y propensas a epidemias a través del desarrollo de estrategias basadas en la evidencia para predecir, prevenir, detectar y responder. **(APOYAR)**
1434. Se han descrito casos de enfermos de COVID que muestran \*problemas neurológicos de alteración del nivel de conciencia\*. **(MOSTRAR)**
1435. A medida que la pandemia iba afectando a más cantidad de países se descubrieron \*nuevos signos del virus que se añaden a los anteriores como por ejemplo, la dificultad para hablar o moverse\*. **(DESCUBRIR)**
1436. A las personas que ya estén infectadas, se les recomienda que se queden \*en casa\*, excepto para recibir atención médica, llamar con antelación antes de visitar a un proveedor de atención médica, usar una mascarilla facial (especialmente en público), tapar la tos y los estornudos con un pañuelo desechable, lavarse las manos regularmente con agua y jabón, y evitar compartir artículos personales del hogar. **(QUEDAR)**
1437. El gobierno federal está proporcionando vacunas contra el COVID-19 autorizadas por la FDA de forma gratuita a todas las personas que viven \*en los Estados Unidos\*, independientemente de su estado migratorio o de seguro médico. **(VIVIR)**
1438. Este gráfico de barras \*nos\* muestra el conteo acumulado de defunciones por COVID-19. **(MOSTRAR)**
1439. \*A finales de enero\*, la Comisión Nacional de Salud de China publicó una guía de manejo de las crisis psicológicas, en la que propugnaba la intervención de las personas afectadas, contactos cercanos, los encerrados en sus hogares, los familiares y amigos de los pacientes, personal sanitario y el público general que lo requiriera. **(PUBLICAR)**
1440. \*Las personas\* deben seguir las directrices establecidas por sus organizaciones locales de salud para garantizar que se mantengan a salvo y mantengan así a los demás. **(SEGUIR)**
1441. De igual modo, en Sudán del Sur, la situación se agravó considerablemente \*durante el confinamiento\*. **(AGRAVAR)**
1442. \*Durante las primeras olas de infección\* se hizo el primer gran estudio sobre los síntomas, y que publicó la prestigiosa The New England Journal of Medicine que acaloraba qué síntomas no eran tan habituales. **(HACER)**
1443. Para ese entonces \*ya\* se habían comunicado los primeros casos por COVID-19 fuera de China: dos en Tailandia y uno en Japón. **(COMUNICAR)**
1444. España hasta el momento suma \*un total de 3.971.124 casos positivos de Covid-19\* y ya son 81.020 las personas fallecidas, según los últimos datos proporcionados por el Ministerio de Sanidad. **(SUMAR)**

1445. Existe evidencia limitada que sugiere \*que el virus podría transmitirse uno o dos días antes de que se tengan síntomas\*, ya que la viremia alcanza un pico al final del período de incubación. **(SUGERIR)**
1446. \*Algunas investigaciones\* han sugerido que COVID-19 es una enfermedad vascular en lugar de respiratoria, lo que significa que puede viajar a través de los vasos sanguíneos. **(SUGERIR)**
1447. Quienes ya se estaban recuperando pueden ver \*que sus fiebres desaparecen\* después de 12 días. **(VER)**
1448. Si tiene COVID-19 y tiene un alto riesgo de enfermedad grave, \*su médico\* puede recomendarle anticuerpos monoclonales. **(RECOMENDAR)**
1449. Los Institutos Nacionales de la Salud (NIH, por sus siglas en inglés) lideran y respaldan la investigación que ayuda a acelerar el desarrollo de pruebas, tratamientos y vacunas para ayudar \*a mejorar y salvar vidas\*. **(AYUDAR)**
1450. Ahora que ya has hecho lo más inmediato, estudia y aplica \*las recomendaciones para el aislamiento domiciliario\* e informa a tus convivientes de que tienen que hacer cuarentena. **(APLICAR)**
1451. \*El espectro clínico de COVID-19\* varía de formas asintomáticas a condiciones clínicas caracterizadas por insuficiencia respiratoria severa que requiere ventilación mecánica y apoyo en una unidad de cuidados intensivos (UCI), a manifestaciones en varios órganos y sistemas produciendo sepsis, shock séptico y síndrome de disfunción multiorgánica. **(VARIAR)**
1452. Especialistas apuntan \*a que esta tendencia se debe a que nuestra sociedad sigue asociando a las mujeres a las labores de cuidados, aunque sea de forma inconsciente\*. **(APUNTAR)**
1453. En otros casos, algunas personas infectadas no desarrollan ningún síntoma, pero pueden contagiar igualmente \*al resto de población\*. **(CONTAGIAR)**
1454. Héctor Raúl Vargas, titular de la División de Medicina Familiar del IMSS, explicó \*que se implementará un código QR que le servirá al paciente saber de inmediato si tiene síntomas de Covid-19\*. **(EXPLICAR)**
1455. Asegúrate \*de que estás al día con información fidedigna\*. **(ASEGURAR)**
1456. Quítese la máscara usando la técnica apropiada (es decir, no toque \*la parte delantera\* sino que quite el encaje por detrás). **(TOCAR)**
1457. El aumento de la utilización de internet ha causado \*un incremento de la violencia digital\*, y en especial de la violencia de género digital. **(CAUSAR)**
1458. Si tiene COVID-19 y tiene un alto riesgo de enfermedad grave, su médico puede recomendar \*le\* anticuerpos monoclonales. **(RECOMENDAR)**
1459. Específicamente, \*182 países\* continuaban la interrupción total de los servicios de educación y 8 lo hacían parcialmente de modo local en distintas jurisdicciones. **(CONTINUAR)**
1460. \*Los estudios realizados en Reino Unido, de seguimiento de síntomas desde los inicios de la pandemia\*, evidenciaron durante la ola de delta que los dolores de cabeza se habían convertido en el síntoma más común de la infección con la variante. **(EVIDENCIAR)**
1461. \*Dermatologic Therapy\*, en diciembre de 2020, reportó que algunos pacientes presentaban cambios en la apariencia de su lengua. **(REPORTAR)**
1462. La Organización Mundial de la Salud (OMS) y la Organización Panamericana de la Salud (OPS) han reconocido \*el estado de pandemia\*, que supone el desarrollo de esta enfermedad. **(RECONOCER)**
1463. Para prevenir la expansión del virus, \*los gobiernos\* han impuesto restricciones de viajes, cuarentenas, confinamientos, aislamiento social, cancelación de eventos, y cierre de establecimientos. **(IMPONER)**
1464. Algunas investigaciones han sugerido que COVID-19 es una enfermedad vascular en lugar de respiratoria, lo que significa \*que\* puede viajar a través de los vasos sanguíneos. **(VIAJAR)**
1465. Los datos han demostrado \*que la sintomatología del coronavirus puede ser leve, similar al de un catarro suave\*. **(DEMOSTRAR)**

1466. En general, el coronavirus afecta menos \*a la parte alta del aparato respiratorio\*, a diferencia del resfriado común aunque, como decimos, las nuevas variantes han cambiado este patrón. **(AFECTAR)**
1467. Los estudios tienen como objetivo determinar si \*los inmunomoduladores\* pueden reducir la necesidad de ventiladores y acortar las estancias hospitalarias. **(REDUCIR)**
1468. ONU destaca \*la labor de la Organización en 2021\* para hacer frente al COVID-19, incluido el programa respaldado por la ONU para ayudar a los países en desarrollo a proteger a sus poblaciones contra el virus, y a las medidas adoptadas para prepararse para futuras crisis sanitarias mundiales. **(DESTACAR)**
1469. Un estudio con 2143 niños sugiere \*que este grupo de la población puede ser un factor crítico en la rápida propagación de la enfermedad\*. **(SUGERIR)**
1470. Limpia y desinfecta \*las canastas para la ropa\* y lávate las manos después. **(DESINFECTAR)**
1471. \*Los ciudadanos\* podrán retirarlos con su Tarjeta Sanitaria o, si pertenecen a alguna mutua y no cuentan con tarjeta sanitaria pública, presentando su DNI o NIE. **(PERTENECER)**
1472. La tos seca, la fiebre y el cansancio fueron los primeros síntomas que se asociaron al coronavirus \*cuando irrumpió en nuestras vidas\*. **(ASOCIAR)**
1473. El Ministerio de Defensa de China asegura haber desarrollado con éxito una vacuna contra el coronavirus y ha autorizado \*las pruebas en humanos\*. **(AUTORIZAR)**
1474. Como el nombre oficial de la enfermedad se estableció después de la creación de este material, \*cualquier mención de nCoV\* se refiere a la COVID-19, la enfermedad infecciosa causada por el coronavirus descubierto más recientemente. **(REFERIR)**
1475. El 2 de febrero de 2020, médicos de Tailandia afirmaron \*haber tratado a un paciente con éxito con una combinación de lopinavir / ritonavir y el medicamento contra la influenza oseltamivir\*. **(AFIRMAR)**
1476. Este mapa dispone de un buscador (recuadro \*arriba a la izquierda\*) para localizar una calle o domicilio; el resultado sitúa la zona que corresponde. **(DISPONER)**
1477. La enfermedad es mortal en raras ocasiones, y hasta ahora las víctimas mortales han sido personas de edad avanzada que \*ya\* padecían una enfermedad crónica como diabetes, asma o hipertensión. **(PADECER)**
1478. Su proveedor puede tomar \*muestras\* con hisopos de la parte posterior de su nariz y garganta. **(TOMAR)**
1479. \*Los científicos\* creen que la variante que nació en India y clasificada por la OMS como Delta, podría estar detrás del cambio en el cuadro clínico. **(CREER)**
1480. En lo que respecta al rumbo ambiental, \*las emisiones de gases de dióxido de nitrógeno\* cayeron a niveles mínimos históricos a 6 % a nivel mundial repartido entre las producciones cesadas en China y la Unión Europea. **(CAER)**
1481. Investigadores chinos descubrieron \*que Arbidol, un medicamento antiviral utilizado para tratar la gripe, podría combinarse con Darunavir, un medicamento empleado en el tratamiento del VIH, para el tratamiento de pacientes con coronavirus\*. **(DESCUBRIR)**
1482. Sin embargo, las observaciones clínicas apuntan a que cuando la repuesta inmune no es capaz de controlar eficazmente el virus, como en personas mayores con un sistema inmune debilitado, el virus se propagaría de forma más eficaz produciendo daño en el tejido pulmonar, lo que activaría \*a los macrófagos y granulocitos\* y conduciría a la liberación masiva de citoquinas proinflamatorias. **(ACTIVAR)**
1483. De acuerdo con los últimos datos del Ministerio de Sanidad, 28.993.301 personas en el país ya se encuentran inmunizadas, lo que representa \*el 61,1% de la población\*. **(REPRESENTAR)**
1484. Las personas que contraen \*el coronavirus\* también pueden tener diarrea, como suele ocurrir en los niños que cogen la gripe. **(CONTRAER)**
1485. Una posible señal de que podrías tener covid-19 es si \*tus síntomas\* no mejoran después de una semana más o menos, pero en realidad empeoran. **(MEJORAR)**

1486. Los Institutos Nacionales de la Salud (NIH, por sus siglas en inglés) lideran y respaldan \*la investigación que\* ayuda a acelerar el desarrollo de pruebas, tratamientos y vacunas para ayudar a mejorar y salvar vidas. **(AYUDAR)**
1487. Una de cada seis personas que contrae el virus desarrolla \*una enfermedad grave\* y tiene gran dificultad para respirar, especialmente personas mayores y las que padecen afecciones médicas subyacentes, como hipertensión arterial, problemas cardíacos o diabetes. **(DESARROLLAR)**
1488. Después de que los investigadores prueban a fondo nuevos tratamientos o vacunas \*en el laboratorio\* para asegurarse de que puedan beneficiar a las personas, los tratamientos más prometedores pasan a los estudios clínicos. **(PROBAR)**
1489. Se han enviado suministros de desinfección, respiradores purificadores de aire, protectores faciales, guantes, reactivos de detección de ácido nucleico de coronavirus, ventiladores, monitores de pacientes, bombas de jeringa, bombas de infusión y alimentos \*a las áreas afectadas\*, e incluso han ofrecido pruebas gratuitas de COVID-19 al público. **(ENVIAR)**
1490. La OMS desaconsejaba en marzo la utilización de máscara quirúrgica por la población sana, \*en abril\* la OMS consideró que era una medida aceptable en algunos países. **(CONSIDERAR)**
1491. \*Dra. Nora Volkow\* habla sobre un estudio que muestra que las personas con trastornos por consumo de sustancias, incluido el trastorno por consumo de marihuana, pueden tener un mayor riesgo de contraer infecciones irruptivas del SARS-CoV-2. **(HABLAR)**
1492. \*Los sabañones\* se forman en respuesta a la exposición repetida al aire frío. **(FORMAR)**
1493. El hallazgo de varias vacunas contra el coronavirus, eficaces para evitar el curso grave de la enfermedad y frenar \*las muertes\*, ha sido el punto de inflexión de la pandemia. **(FREJAR)**
1494. En cambio, los respiradores certificados también dan \*una gran protección\* a su portador. **(DAR)**
1495. La mayoría de los casos de contagio se dan entre la población que tiene de 30 a 79 años, con una tasa de letalidad media del 2,3%, mientras que la mayor parte de los fallecidos supera \*la barrera de los 50 años\*, una edad a la que, según se deduce del mayor estudio elaborado hasta la fecha, aumenta el riesgo de mortalidad, con una tasa de letalidad del 1,3%. **(SUPERAR)**
1496. \*Al principio\*, la ciencia no pensaba que la diarrea u otros problemas gástricos típicos que a menudo vienen con la gripe se aplicarían al nuevo coronavirus, también conocido como SARS-CoV-2. **(PENSAR)**
1497. El proyecto Zoe también demostró que \*las personas que ya han sido vacunadas y dieron positivo\* estornudaban con más frecuencia. **(ESTORNUDAR)**
1498. Si bien las medidas de contención previenen \*la propagación del virus\*, las mujeres y niñas se ven obligadas a convivir con personas que son violentas y quedan aisladas de los recursos que pueden ayudarlas. **(PREVENIR)**
1499. \*La publicación\* recuerda que el Servicio Nacional de Salud (NHS) británico actualmente solo menciona tres síntomas oficiales de covid: la fiebre, tos y cambios en el gusto o el olfato. **(RECORDAR)**
1500. Hay \*mascarillas que\* llevan alguna válvula de exhalación (es decir, de salida), la cual expulsa el aire de la respiración sin filtrar. **(LLEVAR)**
1501. Sus investigadores han analizado \*los síntomas vinculados a la variante delta y la nueva variante ómicron\*. **(ANALIZAR)**
1502. Además, \*hay otros países como España y Bélgica que\* registran un número desigual entre hombres y mujeres durante toda la pandemia. **(REGISTRAR)**
1503. \*La pandemia\* también ha afectado al entorno de trabajo de las mujeres dominicanas. **(AFECTAR)**
1504. \*La India\* registra casi 200.000 nuevos casos de coronavirus en las últimas 24 horas **(REGISTRAR)**
1505. En una encuesta, \*más del 50% de las mujeres con responsabilidades familiares\* respondieron que se encargaban ellas solas de la enseñanza de sus hijos o de ayudar a algún familiar con dependencia, mientras que en el caso de los hombres solo el 8% afirmaban que en su casa ocurriera este hecho. **(RESPONDER)**

1506. \*Algunos virus aéreos\* pueden acabar en el tracto digestivo y provocar problemas digestivos. **(ACABAR)**
1507. \*Downing Street\* pide disculpas a Isabel II por las fiestas celebradas en la víspera del funeral de su marido. **(PEDIR)**
1508. Las vacunas protegen \*a las personas\* de infectarse o de enfermarse gravemente, y las mascarillas y el distanciamiento social ayudan a evitar que el virus se propague**(PROTEGER)**
1509. En una encuesta, más del 50% de las mujeres con responsabilidades familiares respondieron \*que se encargaban ellas solas de la enseñanza de sus hijos o de ayudar a algún familiar con dependencia\*, mientras que en el caso de los hombres solo el 8% afirmaban que en su casa ocurriera este hecho. **(RESPONDER)**
1510. Además, hay otros países como España y Bélgica que registran un número desigual entre hombres y mujeres \*durante toda la pandemia\*. **(REGISTRAR)**
1511. En general, creo que estamos obteniendo \*un poco más de información sobre los tipos de síntomas que los pacientes pueden tener\*, dijo Gupta. **(OBTENER)**
1512. Álvarez, que es miembro del Comité Asesor Epidemiológico del Ministerio de Salud, explicó \*que los síntomas más comunes de la variante ómicron son los asociados con síntomas respiratorios altos, especialmente dolor de cabeza, producción de mocos, dolor de garganta, tos seca\*. **(EXPLICAR)**
1513. \*El Oso y Madroño de la artista Noelia Jiménez, teñido de una amplia paleta de colores vivos\*, representa cómo los trabajadores esenciales han unido a la sociedad española durante la pandemia. **(REPRESENTAR)**
1514. \*Otros\* han registrado la pérdida súbita del olfato y el gusto de forma menos frecuente (sin que la mucosidad fuese la causa). **(REGISTRAR)**
1515. La transmisión del SARS-CoV-2 se produce mediante pequeñas gotas –microgotas de Flügge– que se emiten al hablar, estornudar, toser o espirar, que al ser despedidas por un portador (que puede no tener síntomas de la enfermedad o estar incubándola) pasan directamente a otra persona mediante la inhalación, o quedan sobre los objetos y superficies que rodean al emisor, y luego, a través de \*las manos, que\* lo recogen del ambiente contaminado, toman contacto con las membranas mucosas orales, nasales y oculares, al tocarse la boca, la nariz o los ojos. **(RECOGER)**
1516. La transmisión del SARS-CoV-2 se produce mediante pequeñas gotas –microgotas de Flügge– que se emiten al hablar, estornudar, toser o espirar, que al ser despedidas por un portador (que puede no tener síntomas de la enfermedad o estar incubándola) pasan directamente a otra persona mediante la inhalación, o quedan sobre los objetos y superficies que rodean al emisor, y luego, a través de las manos, que \*lo\* recogen del ambiente contaminado, toman contacto con las membranas mucosas orales, nasales y oculares, al tocarse la boca, la nariz o los ojos. **(RECOGER)**
1517. España supera los ocho millones de casos registrados, tras sumar \*casi un millón\* en una semana. **(SUMAR)**
1518. No obstante, \*este tiempo de permanencia\* depende de factores como la temperatura y el tipo de superficie. **(DEPENDER)**
1519. Si una persona obtiene un resultado positivo en COVID-19, debe llamar al teléfono 900 102 112 para que le recojan \*sus datos\* y le informen de las medidas que debe adoptar. **(RECOGER)**
1520. La pandemia también ha afectado \*al entorno de trabajo de las mujeres dominicanas\*. **(AFECTAR)**
1521. Aún no hemos llegado a ese punto, ya que todavía tenemos que probarlo en ensayos clínicos, pero es el objetivo final, afirmó \*Michael Snyder\*. **(AFIRMAR)**
1522. Realiza el lavado de manos frecuentemente o bien, desinfectar \*las\* con gel con base de alcohol al 70%**(DESINFECTAR)**
1523. Aunque \*el virus\* ataca primero a los pulmones, también puede infectar el corazón, los riñones, el hígado, el cerebro y los intestinos. **(INFECTAR)**
1524. Los estudios también indican que la protección contra el SARS-CoV-2 comienza a disminuir con el tiempo \*después de las dosis iniciales de vacuna\*. **(DISMINUIR)**

1525. La gente vacunada todavía puede transmitir \*la COVID\*, aunque es más difícil. **(TRANSMITIR)**
1526. \*Las máscaras\* limitan el volumen y la distancia de viaje de las gotas respiratorias dispersas al hablar, estornudar y toser. **(LIMITAR)**
1527. En los últimos días, ante el aumento de contagios por variante Ómicron del COVID-19 y en medio de la recién iniciada tercera ola en el Perú, se ha vuelto recurrente \*la presencia de síntomas de SarCov-2 en la población\*. **(VOLVER)**
1528. Los datos analizados por la app mostraron que la tos era el síntoma más común al comienzo de la pandemia, y el 46 por ciento de los pacientes infectados padecían \*este signo revelador\*. **(PADECER)**
1529. FDA aprobó \*una vacuna contra el COVID-19\* y dos fueron autorizadas para su uso de emergencia. **(APROBAR)**
1530. \*Isaac González, publicista\*, recuerda así la vida que vuelve y la que se va, especialmente la de los ancianos, los más golpeados por la covid **(RECORDAR)**
1531. La enfermedad se propaga principalmente de persona a persona a través de las gotículas que salen despedidas de la nariz o la boca de una persona infectada \*al toser, estornudar o hablar\*. **(SALIR)**

## Anexo C

### Lista de verbos seleccionados

ACABAR
ACONSEJAR
ACTIVAR
ACTUALIZAR
ACTUAR
ACUDIR
ADMINISTRAR
ADVERTIR
AFECTAR
AFIRMAR
AGRAVAR
AGREGAR
ALCANZAR
ANALIZAR
ANUNCIAR
APARECER
APLICAR
APOYAR
APROBAR
APUNTAR
ASEGURAR
ASOCIAR
ATENDER
AUMENTAR
AUTORIZAR
AVANZAR
AYUDAR
BUSCAR
CAER
CAMBIAR
CANCELAR
CAUSAR
CERRAR
COINCIDIR
COMENZAR
COMER

COMPARAR
COMPARTIR
COMPLETAR
COMUNICAR
CONFIRMAR
CONOCER
CONSIDERAR
CONSULTAR
CONTAGIAR
CONTAR
CONTENER
CONTINUAR
CONTRAER
CONTROLAR
CONVERTIR
CREAR
CREER
CUBRIR
CUIDAR
DAR
DEJAR
DEMOSTRAR
DEPENDER
DESAPARECER
DESARROLLAR
DESCARTAR
DESCONOCER
DESCRIBIR
DESCUBRIR
DESECHAR
DESINFECTAR
DESTACAR
DETECTAR
DETERMINAR
DIAGNOSTICAR
DIRIGIR

DISMINUIR
DISPONER
DURAR
ELEVAR
ELIMINAR
EMPEORAR
EMPEZAR
ENCONTRAR
ENFERMAR
ENTRAR
ENVIAR
ESCRIBIR
ESPERAR
ESTABLECER
ESTIMAR
ESTORNUDAR
ESTUDIAR
EVALUAR
EVIDENCIAR
EVITAR
EXCLUIR
EXISTIR
EXPERIMENTAR
EXPLICAR
EXTENDERSE
FALLECER
FORMAR
FRENAR
FUNCIONAR
GENERAR
HABER
HABLAR
HACER
IDENTIFICAR
IMPONER
INCLUIR

INCREMENTAR
INDICAR
INFECTAR
INFORMAR
INGRESAR
INVESTIGAR
IR
LAVAR
LIMITAR
LIMPIAR
LLAMAR
LLEGAR
LLEVAR
LOGRAR
MANIFESTAR
MANTENER
MEJORAR
MORIR
MOSTRAR
NECESITAR
NOTIFICAR
OBSERVAR
OBTENER
OCURRIR
OFRECER
PADECER
PASAR
PEDIR
PENSAR
PERDER
PERMANECER

PERMITIR
PERTENECER
PONER
PRESENTAR
PREVENIR
PROBAR
PRODUCIR
PROPAGAR
PROPORCIONAR
PROTEGER
PROVOCAR
PUBLICAR
QUEDAR
QUERER
REALIZAR
RECIBIR
RECOGER
RECOMENDAR
RECONOCER
RECORDAR
RECUPERAR
REDUCIR
REFERIR
REGISTRAR
REPORTAR
REPRESENTAR
REQUERIR
RESPIRAR
RESPONDER
RESULTAR
REVISAR

SABER
SALIR
SEGUIR
SENTIR
SEÑALAR
SEPARAR
SIGNIFICAR
SUCEDER
SUFRIR
SUGERIR
SUMAR
SUPERAR
SUPONER
SURGIR
TARDAR
TENER
TERMINAR
TOCAR
TOMAR
TOSER
TRABAJAR
TRANSMITIR
TRATAR
USAR
UTILIZAR
VARIAR
VER
VIAJAR
VISITAR
VIVIR
VOLVER