

# Tarea 1 del Taller NEGES 2018: Guías de Anotación

## *NEGES 2018 Task 1: Annotation Guidelines*

Salud María Jiménez-Zafra<sup>1</sup>, Noa P. Cruz Díaz<sup>2</sup>,  
Roser Morante<sup>3</sup>, María Teresa Martín-Valdivia<sup>1</sup>

<sup>1</sup>SINAI, Centro de Estudios Avanzados en TIC (CEATIC), Universidad de Jaén

{sjzafra, maite}@ujaen.es

<sup>2</sup>Savana Médica, Madrid

contact@noacruz.com

<sup>3</sup>CLTL Lab, Computational Linguistics, VU University Amsterdam

r.morantevallejo@vu.nl

**Resumen:** En este trabajo se presenta la tarea 1 del taller NEGES 2018, taller de NEGación en ESpañol, celebrado como parte de la XXXIV edición del Congreso Internacional de la Sociedad Española para el Procesamiento del Lenguaje Natural. El objetivo de esta tarea es llegar a un acuerdo sobre las directrices a seguir para la anotación de la negación en textos escritos en español con el fin de poder desarrollar sistemas capaces de procesar este fenómeno.

**Palabras clave:** Negación, directrices de anotación, NEGES 2018, SEPLN

**Abstract:** This paper presents task 1 of NEGES 2018, Workshop on Negation in Spanish, held as part of the XXXIV edition of the International Conference of the Spanish Society for Natural Language Processing. The objective of this task is to reach an agreement on the guidelines to follow for the annotation of negation in Spanish texts in order to develop negation processing systems.

**Keywords:** Negation, annotation guidelines, NEGES 2018, SEPLN

## 1 Introducción

La negación es un elemento lingüístico clave en muchas aplicaciones relacionadas con el Procesamiento del Lenguaje Natural (PLN), como el análisis de sentimientos (Liu, 2012) o la recuperación de información (Büttcher, Clarke, y Cormack, 2010), ya que su presencia en una oración puede cambiar totalmente su significado (Horn, 1989). En un sistema de análisis de opiniones, la valoración de una película debería ser totalmente diferente si los espectadores dicen “La película me gustó mucho” que si dicen “La película no me gustó nada”. Un sistema de recuperación de información como Google debería devolver diferentes resultados ante las consultas “películas que no sean de aventuras” y “películas que sean de aventuras”, pero no es así. Al escuchar estas frases, las personas somos capaces de diferenciar el significado de cada una de ellas porque tenemos asimilado el concepto de negación. Si queremos que un sistema sea capaz de ello, es necesario enseñarle cómo detectar la presencia de marcadores de negación en una oración y cómo identificar su ámbito

de influencia, es decir, cuáles son las palabras cuyo significado se está viendo modificado.

Hasta ahora, la mayor parte de la investigación sobre la negación se ha realizado sobre textos en inglés. Sin embargo, la presencia de otros idiomas en Internet, entre los que se encuentra el español, es cada vez mayor. Por ello, este taller tiene como objetivo avanzar en el estudio de la negación en español, ya que es una lengua bastante compleja en la que aparte de los mecanismos habituales de negación, se utiliza también la doble negación.

La negación no sólo se emplea para decir NO, como podemos pensar habitualmente. En ocasiones utilizamos la negación para reforzar el valor positivo de algo. Por ejemplo, en la oración “No puede ser más bonito” estamos haciendo uso de la clave de negación “No” para reforzar el valor positivo del adjetivo “bonito”.

En español, también encontramos un uso habitual de la doble negación. El lenguaje lógico es muy claro y estricto, dos negaciones forman una afirmación, pero el lenguaje natural no siempre es uniforme. Podemos en-

contrar casos en los que dos negaciones tienen un sentido positivo, por ejemplo, “No es verdad que no me apasionen las matemáticas”, o un sentido negativo, “No sé nada de astrología”. En el primer caso, se niega que no me apasionen las matemáticas, lo cual quiere decir que me apasionan. Mientras que en el segundo caso, “nada” está reforzando el sentido negativo del enunciado.

La negación es un elemento del lenguaje bastante complejo para el que la disponibilidad de corpus anotados es fundamental porque nos va a permitir entrenar sistemas para que sean capaces de procesar de forma adecuada este fenómeno. Sin embargo, no sólo es importante que los corpus estén disponibles sino que también deben estar bien documentados y tener anotaciones de calidad. Además, para poder entrenar un sistema de forma adecuada las anotaciones deben cubrir el mayor número posible de casos y seguir unas directrices comunes. Por ello, en el taller NEGES 2018<sup>1</sup> se ha propuesto como tarea 1 una revisión de las guías de anotación existentes en español, con el objetivo de poder definir unas directrices comunes.

El resto del documento está organizado como sigue: en la Sección 2 se describe la tarea 1 del taller NEGES 2018, en la Sección 3 se presentan los corpus cuyas directrices de anotación se han analizado, en la Sección 4 se presentan los participantes de la tarea y el análisis realizado y, por último, en la Sección 5 se presentan las conclusiones.

## 2 Descripción de la tarea

La tarea 1 del taller NEGES 2018, “Guías de anotación”, tiene como objetivo llegar a un acuerdo sobre las directrices a seguir para anotar la negación en textos escritos en español. Jiménez-Zafra et al. (2018a) realizaron una recopilación de los corpus anotados con negación en español hasta el momento, mostrando que no son compatibles por diferencias en el género, en las directrices de anotación empleadas, en los aspectos de la negación anotados (clave de negación, alcance, evento, foco) y en los patrones de negación considerados (Martí et al., 2016). Para poder desarrollar un sistema automático capaz de procesar la negación es fundamental disponer de corpus anotados con esta información, que cubran el mayor número posible

de casos y que tengan un tamaño considerable. Los corpus anotados hasta ahora pertenecen a 3 dominios (noticias, informes clínicos y opiniones de productos) y se basan en directrices diferentes, por lo que no se pueden unir para entrenar un sistema de aprendizaje automático.

A pesar de que ya ha habido varios esfuerzos de anotación, la comunidad carece de un estándar para anotar la negación, al contrario de lo que sucede con otros fenómenos, como los roles semánticos. La tarea 1 del taller NEGES 2018 se propuso con este fin, tratar de llegar a un acuerdo para generar un estándar para la anotación de la negación en español. Para ello, las guías utilizadas para la anotación de los corpus existentes hasta el momento se han puesto a disposición de los participantes para que puedan analizarlas. Se ha proporcionado un período de análisis y una vez transcurrido el mismo, los participantes han enviado un documento indicando con qué aspectos de las guías están de acuerdo y con cuáles no, todo ello debidamente justificado. Esta información se ha enviado a los participantes, previamente a la celebración del taller, para que en el transcurso del taller se puedan discutir los aspectos de interés y llegar a un consenso.

## 3 Corpus analizados

En la tarea 1 se han analizado los 5 corpus en español anotados con negación hasta el momento, que pertenecen a 3 dominios diferentes: noticias, informes clínicos y opiniones de productos. A continuación, se describe brevemente cada uno de ellos.

### 3.1 UAM Spanish Treebank

Sandoval y Salazar (2013) fueron los primeros investigadores que enriquecieron un corpus en español, el UAM Spanish Treebank (Moreno et al., 2003), con anotaciones sobre claves de negación y sus respectivos ámbitos.

El corpus UAM Spanish Treebank está formado por 1.500 oraciones extraídas de artículos periodísticos de *El País Digital* y *Compra Maestra* que inicialmente fueron anotadas sintácticamente por Moreno et al. (2003). Posteriormente, 10 años más tarde, Sandoval y Salazar (2013) enriquecieron esta versión del corpus con la anotación de claves de negación sintácticas y con sus respectivos ámbitos, identificando que el 10.67 % de las oraciones contenían negación, un total de 160

<sup>1</sup><http://www.sepln.org/workshops/neges/index.php>

oraciones.

Las anotaciones fueron realizadas por dos expertos en lingüística de corpus que siguieron unas directrices similares a las utilizadas para la anotación del corpus BioScope (Vincze et al., 2008). Las pautas empleadas para la anotación del corpus se encuentran en las páginas 51-55 del artículo de descripción del corpus (Sandoval y Salazar, 2013) y es la información que fue revisada por los participantes de la tarea.

### 3.2 IxaMed-GS

El corpus IxaMed-GS (Oronoz et al., 2015) está formado por 75 registros electrónicos de salud del Hospital Galdakao-Usansolo de Vizcaya. Fue anotado por dos expertos en farmacología y farmacovigilancia con entidades relacionadas con enfermedades y medicamentos, y con las relaciones entre ellas que indican efectos adversos. El principal objetivo de la anotación fue la identificación de entidades y eventos, pero la negación y la especulación se tuvieron en cuenta en el proceso de anotación. De esta manera, se anotaron 2.362 enfermedades, de las cuales 490 (20,75 %) se etiquetaron como enfermedades negadas y 40 (1,69 %) como enfermedades especuladas. Además, se identificaron 404 entidades relacionadas con alergias, de las cuales 273 (67,57 %) fueron negadas y 13 (3,22 %), especuladas.

Para la anotación del corpus, los anotadores definieron sus propias pautas de anotación teniendo en cuenta las consideraciones de Ananiadou y McNaught (2006). Estas directrices son las que fueron analizadas por los participantes de la tarea y se encuentran en la página 322 del artículo de descripción del corpus (Oronoz et al., 2015).

### 3.3 SFU Review<sub>SP</sub>-NEG

El corpus SFU Review<sub>SP</sub>-NEG<sup>2</sup> (Jiménez-Zafra et al., 2018b) es el primer corpus en español en el que se realiza la anotación de marcadores de negación discontinuos y en el que se define una tipología de patrones de negación para llevar a cabo la anotación (Martí et al., 2016). Además, es el primer corpus donde se anota el efecto de la negación sobre las palabras que están dentro de su alcance, es decir, si la negación provoca un cambio en la

polaridad, un incremento o una reducción de su valor.

Este corpus es una versión enriquecida de la parte en español del corpus SFU de Taboada, Anthony, y Voll (2006). Se trata de un corpus formado por 400 comentarios extraídos de la web *Ciao.es* que pertenecen a 8 dominios diferentes: automóviles, hoteles, lavadoras, libros, teléfonos, música, ordenadores y películas. La versión inicial del corpus (Taboada, Anthony, y Voll, 2006) contiene información sobre la polaridad de los comentarios en base al número de estrellas otorgado por la persona que ha escrito el comentario<sup>3</sup>, teniendo un total de 25 comentarios positivos y 25 comentarios negativos por cada dominio. La versión enriquecida del corpus, denominada SFU Review<sub>SP</sub>-NEG (Jiménez-Zafra et al., 2018b), ha sido anotada de forma automática a nivel de token con la categoría gramatical y el lema, y de forma manual a nivel de oración con marcadores de negación sintácticos y sus correspondientes ámbitos y eventos. En total, de las 9.455 oraciones que forman el corpus, el 31.97 % contienen al menos un marcador de negación (3.022 oraciones). Además, como se ha indicado inicialmente en la descripción del corpus, cada estructura de negación ha sido anotada con el tipo del patrón de negación empleado y con el efecto que provoca la negación sobre las palabras de su alcance.

Las anotaciones fueron realizadas por dos investigadores experimentados con amplia experiencia en la anotación de corpus que supervisaron todo el proceso y dos anotadores entrenados que llevaron a cabo la tarea de anotación. Se tuvieron en cuenta las directrices utilizadas para la anotación del corpus Bioscope (Vincze et al., 2008), pero después de un análisis exhaustivo de la negación en español, se definió una tipología de patrones de negación en español (Martí et al., 2016). Las pautas utilizadas para la anotación de este corpus, que fueron analizadas por los participantes, se encuentran en las páginas 538-559 del artículo de descripción del corpus (Jiménez-Zafra et al., 2018b).

### 3.4 UHU-HUVR

El corpus UHU-HUVR (Cruz et al., 2017) es el primer corpus en español en el que se rea-

<sup>2</sup>Fue publicado online el 22 de mayo de 2017, por eso se sitúa cronológicamente en tercer lugar <https://doi.org/10.1007/s10579-017-9391-x>

<sup>3</sup>1-2 estrellas = comentario negativo, 3 estrellas no se tuvieron en cuenta, y 4-5 estrellas = comentario positivo

liza la anotación de la negación morfológica, también conocida como negación afijal. Está compuesto por 604 informes clínicos del Hospital Virgen del Rocío de Sevilla, de los cuales 276 se corresponden con informes de radiología y 328 con la historia personal de informes de anamnesis escritos en texto libre. Fue anotado con marcadores de negación sintácticos, léxicos y morfológicos y con los eventos negados. En total, de las 3.065 oraciones de informes de anamnesis, el 35,20 % contienen claves de negación (1.079 oraciones). En los informes de radiología, este porcentaje es menor, de 5.347 oraciones, el 22,80 % contienen negación (1.219 oraciones).

La anotación fue realizada por dos anotadores expertos en el dominio teniendo en cuenta las directrices seguidas para la anotación del corpus Thyme (Styler IV et al., 2014) con algunas adaptaciones. Los detalles de las pautas empleadas para la anotación del corpus que los participantes tenían que revisar se encuentran en las páginas 54-57 del artículo de descripción del corpus (Cruz et al., 2017).

### 3.5 IULA Spanish Clinical Record

El último corpus en español anotado con negación del que tenemos consciencia es el IULA Spanish Clinical Record (Marimon et al., 2017). Se trata de un corpus formado por 300 informes clínicos anónimos de varios servicios de uno de los principales hospitales de Barcelona. Contiene anotaciones sobre marcadores de negación sintácticos y léxicos y sobre sus respectivos ámbitos. En total contiene 3.194 oraciones, de las cuales el 34,22 % fueron anotadas con claves de negación (1.093 oraciones).

La anotación del corpus fue realizada por tres lingüistas computacionales asesorados por un clínico. Los anotadores definieron sus propias pautas de anotación teniendo en cuenta las directrices de Mutalik, Deshpande, y Nadkarni (2001) y las empleadas para la anotación del corpus BioScope (Szarvas et al., 2008) y del corpus ConanDoyle-neg (Morante y Daelemans, 2012). Estas pautas son las que los participantes de la tarea 1 tenían que analizar y se encuentran definidas en las páginas 45-49 del artículo de descripción del corpus (Marimon et al., 2017).

## 4 Participantes y análisis

La tarea 1 de NEGES 2018 ha tenido dos participantes: el equipo CLiC formado por M.

Antonia Martí y Mariona Taulé de la Universidad de Barcelona, y Lucia Donatelli de la Universidad de Georgetown.

Martí y Taulé (2018) comienzan el análisis comparando el *tipo de información* de los corpus. Los 3 corpus del dominio biomédico (IxaMed-GS, UHU-HUVR e IULA Spanish Clinical Record) contienen información semi-estructurada, mientras que los otros dos corpus (SFU Review<sub>sp</sub>-NEG y UAM Spanish Treebank) contienen información no estructurada. Posteriormente se centran en analizar cinco aspectos de relevancia para la anotación de la negación: i) el marcador de negación, ii) el ámbito y la inclusión del sujeto en el ámbito, iii) las estructuras coordinadas, iv) las locuciones de negación y v) la negación léxica y afijal. A continuación resumimos el análisis realizado por Martí y Taulé sobre cada uno de los aspectos:

1. *Marcador de negación.* Las autoras destacan el hecho de que en todos los corpus se anota el marcador de negación, excepto en el corpus IxaMed-GS, e indican que se debe a que el objetivo del mismo sólo es extraer qué entidades están negadas y, por lo tanto, ignoran el marcador.
2. *Ámbito.* En relación al ámbito, todos los corpus lo han anotado, pero divergen en si incluyen (SFU Review<sub>sp</sub>-NEG y UAM Spanish Treebank) o no (UHU-HUVR e IULA Spanish Clinical Record) el sujeto dentro del ámbito. Indican que este hecho puede deberse a las características del dominio biomédico, ya que muchas de las expresiones utilizadas no tienen estructura oracional (1), por lo que no tiene sentido marcar el sujeto.
3. *Estructuras coordinadas.* Respecto a la negación en estructuras coordinadas, también observan diferencias. Sólo se tratan en los corpus SFU Review<sub>sp</sub>-NEG, UHU-HUVR e IULA Spanish Clinical Record, aunque de manera distinta. En el corpus SFU Review<sub>sp</sub>-NEG se distingue entre las estructuras negativas coordinadas (2), donde cada marcador de negación es independiente y tiene su propio ámbito, y las estructuras negativas que contienen marcadores de negación discontinuos (3), donde el ámbito incluye toda la coordinación. Sin embargo, en el corpus IULA Spanish Clinical

Record siempre se incluye la coordinación dentro del ámbito (4), mientras que en el corpus UHU-HUVR cualquier tipo de coordinación da lugar a dos estructuras distintas con sus marcadores de negación y correspondientes ámbitos (5) y (6).

4. *Locuciones de negación.* Con respecto a las locuciones de negación, expresiones multipalabra que expresan negación (ej. en la vida), sólo se tratan en el corpus SFU Review<sub>SP</sub>-NEG.
5. *Negación léxica y afijal.* Por último, en cuanto a la negación léxica y afijal, sólo se anota en los corpus UHU-HUVR e IULA Spanish Clinical Record, aunque de forma limitada. IULA Spanish Clinical Record limita la negación léxica a los predicados ‘descartar’, ‘ausencia de’ e ‘incapaz de’, y en el corpus UHU-HUVR, sólo se mencionan los negadores ‘abandono’, ‘negativo’, ‘-’ y ‘/’, pero no se dan más detalles.

(1) No alergias medicamentosas  
(UHU-HUVR)

(2) No [soy muy alta] tampoco [un pitufo]  
(SFU Review<sub>SP</sub>-NEG)

(3) No [es ni muy pesado ni muy ligero]  
(SFU Review<sub>SP</sub>-NEG)

(4) No [masas ni megalias]  
(IULA Spanish Clinical Record)

(5) No hemos observado [alteraciones a nivel de los distintos ligamentos (...)], así como [de las restantes partes blandas]  
(UHU-HUVR)

(6) No [hay evidencia de módulos pulmonares] / [adenomegalias mediastínicas]  
(UHU-HUVR)

En cuanto a las directrices propuestas para tratar de llegar a un consenso sobre la anotación de la negación en español, Martí y Taulé (2018) proponen lo siguiente:

1. Anotar los marcadores de negación, distinguiendo entre marcadores simples (ej. no, sin) y complejos (ej. no...nadie).
2. Anotar el ámbito de la negación incluyendo el sujeto dentro del mismo.

3. Realizar un tratamiento de la negación coordinada.

4. Anotar las locuciones negativas (ej. en absoluto), aunque no contengan marcadores explícitos de negación.

5. Anotar la negación léxica y morfológica.

6. Anotar el foco de la negación, que no se trata en ninguna de las guías analizadas.

Donatelli (2018) describe de forma individual cada corpus e indica los elementos que echa en falta en la anotación de cada uno de ellos y aquellos aspectos que deberían haberse tenido en cuenta. Considera que se pueden combinar algunos componentes de las propuestas de anotación de los distintos corpus para dar lugar a un conjunto de directrices lingüísticamente precisas y neutras en cuanto al dominio. Indica que para representar la semántica de la negación se deben anotar los siguientes elementos:

1. La clave de negación: elemento léxico que expresa la negación.

2. El ámbito: parte del texto que es negada.

3. El foco: parte del ámbito que es negada de forma prominente o explícita.

4. El refuerzo (en caso de que exista): negación auxiliar o elemento de polaridad negativa, conocido como NPI (Negative Polarity Item) (Altuna, Minard, y Speranza, 2017).

A continuación podemos ver, en un ejemplo proporcionado por la autora (7), los diferentes elementos explicados anteriormente. La clave de negación aparece en negrita, el ámbito en cursiva, el foco entre corchetes y el refuerzo en negrita y en cursiva.

(7) Juan **no** come [carne] **sino** verduras.

Donatelli considera que el esquema propuesto por Jiménez-Zafra et al. (2018b) para la anotación del corpus SFU Review<sub>SP</sub>-NEG es adecuado para capturar las capas de complejidad de la negación y propone combinarlo con el uso de la etiqueta *NegPolItem* utilizada por Marimon et al. (2017) en la anotación del corpus IULA Spanish Clinical Record para marcar los ítems de polaridad negativa (NPI) o negaciones auxiliares.

## 5 Conclusiones

La negación es un fenómeno lingüístico complejo que se ha estudiado ampliamente desde una perspectiva teórica. La detección de aserciones negativas es relevante en una amplia gama de aplicaciones, como la extracción de información, la traducción automática o el análisis de sentimientos, donde es crucial conocer cuándo una parte del texto tiene un significado diferente por la presencia de la negación. La detección de la negación se está convirtiendo en una tarea relevante para el PLN, de hecho, algunos autores lo definen como un desafío (Liu, 2012). Sin embargo, la mayor parte de la investigación sobre la negación se ha realizado para el inglés. Por lo tanto, con este taller, se pretende avanzar en el estudio de este fenómeno en español, la segunda lengua más hablada en el mundo y la tercera más utilizada en Internet.

El principal objetivo de la tarea 1 es reunir a la comunidad científica que está trabajando en este tema para poder llegar a un acuerdo sobre las directrices a seguir para anotar la negación, ya que aunque ha habido varios esfuerzos de anotación, la comunidad carece de un estándar, lo que dificulta el desarrollo de sistemas automáticos para procesar este fenómeno.

El análisis realizado por los participantes se utilizará como punto de partida para que en el transcurso del taller se puedan discutir los aspectos de interés y llegar a un consenso. Asimismo, la discusión originada marcará las líneas de futuro de esta tarea.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Educación, Cultura y Deporte (MECD - ayuda FPU014/00983), el Fondo Europeo de Desarrollo Regional (FEDER) y el proyecto REDES (TIN2015-65136-C2-1-R) del Gobierno de España. Roser Morante ha sido financiada por la Organización Holandesa para la Investigación Científica (NWO) a través del premio Spinoza otorgado a Piek Vossen (SPI 30-673, 2014-2019).

## Bibliografía

- Altuna, B., A.-L. Minard, y M. Speranza. 2017. The Scope and Focus of Negation: A Complete Annotation Framework for Italian. En *Proceedings of the Workshop Computational Semantics Beyond Events and Roles*, páginas 34–42.
- Ananiadou, S. y J. McNaught. 2006. *Text mining for biology and biomedicine*. Artech House London.
- Büttcher, S., C. L. Clarke, y G. V. Cormack. 2010. *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press.
- Cruz, N., R. Morante, M. J. M. López, J. M. Vázquez, y C. L. P. Calderón. 2017. Annotating negation in Spanish clinical texts. En *Proceedings of the Workshop Computational Semantics Beyond Events and Roles*, páginas 53–58.
- Donatelli, L. 2018. Cues, Scope, and Focus: Annotating Negation in Spanish Corpora. En *Proceedings of NEGES 2018: Workshop on Negation in Spanish*, volumen 2174, páginas 29–34.
- Horn, L. R. 1989. *A natural history of negation*. CSLI Publications.
- Jiménez-Zafra, S. M., R. Morante, M. Martín, y L. A. U. Lopez. 2018a. A review of Spanish corpora annotated with negation. En *Proceedings of the 27th International Conference on Computational Linguistics*, páginas 915–924.
- Jiménez-Zafra, S. M., M. Taulé, M. T. Martín-Valdivia, L. A. Ureña-López, y M. A. Martí. 2018b. SFU ReviewSP-NEG: a Spanish corpus annotated with negation for sentiment analysis. A typology of negation patterns. *Language Resources and Evaluation*, 52(2):533–569.
- Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.
- Marimon, M., J. Vivaldi, N. Bel, y R. Boronat. 2017. Annotation of negation in the IULA Spanish Clinical Record Corpus. *SemBEaR 2017*, 5(36.41):43.
- Martí, M. A. y M. Taulé. 2018. Análisis Comparativo de los Sistemas de Anotación de la Negación en Español. En *Proceedings of NEGES 2018: Workshop on Negation in Spanish*, volumen 2174, páginas 23–28.
- Martí, M. A., M. Taulé, M. Nofre, L. Marsó, M. T. Martín-Valdivia, y S. M. Jiménez-Zafra. 2016. La negación en español: análisis y tipología de patrones de negación. *Procesamiento del Lenguaje Natural*, (57):41–48.

- Morante, R. y W. Daelemans. 2012. ConanDoyle-neg: Annotation of negation in Conan Doyle stories. En *Proceedings of the Eighth International Conference on Language Resources and Evaluation, Istanbul*. Citeseer.
- Moreno, A., S. López, F. Sánchez, y R. Grishman. 2003. Developing a syntactic annotation scheme and tools for a Spanish treebank. En *Treebanks*. Springer, páginas 149–163.
- Mutalik, P. G., A. Deshpande, y P. M. Nadkarni. 2001. Use of general-purpose negation detection to augment concept indexing of medical documents: a quantitative study using the UMLS. *Journal of the American Medical Informatics Association*, 8(6):598–609.
- Oronoz, M., K. Gojenola, A. Pérez, A. D'íaz de Ilarraza, y A. Casillas. 2015. On the creation of a clinical gold standard corpus in Spanish: Mining adverse drug reactions. *Journal of biomedical informatics*, 56:318–332.
- Sandoval, A. M. y M. G. Salazar. 2013. La anotación de la negación en un corpus escrito etiquetado sintácticamente. Annotation of negation in a written treebank. *Revista Iberoamericana de Linguística*, 8.
- Styler IV, W. F., S. Bethard, S. Finan, M. Palmer, S. Pradhan, P. de Groen, B. Erickson, T. Miller, C. Lin, G. Savova, y others. 2014. Temporal annotation in the clinical domain. *Transactions of the Association for Computational Linguistics*, 2:143–154.
- Szarvas, G., V. Vincze, R. Farkas, y J. Csirik. 2008. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. En *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, páginas 38–45. Association for Computational Linguistics.
- Taboada, M., C. Anthony, y K. Voll. 2006. Methods for creating semantic orientation dictionaries. En *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06)*, páginas 427–432.
- Vincze, V., G. Szarvas, R. Farkas, G. Móra, y J. Csirik. 2008. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC bioinformatics*, 9(11):1.

