

El cambio de polaridad en la minería de opiniones a través de la cuantificación en inglés

Polarity shifting in opinion mining through quantification in English

Yolanda Blázquez-López

Universidad Nacional de Educación a Distancia (UNED)
España

Carlos Perriñán-Pascual

Universitat Politècnica de València
España

ONOMÁZEIN 65 (septiembre de 2024): 206-226

DOI: 10.7764/onomazein.65.12

ISSN: 0718-5758



Yolanda Blázquez-López: Facultad de Filología, Universidad Nacional de Educación a Distancia (UNED), España.
| E-mail: yblazquez6@alumno.uned.es

Carlos Perriñán-Pascual: Departamento de Lingüística Aplicada, Universitat Politècnica de València, España.
Orcid: 0000-0002-6483-4712. | E-mail: joepas3@upv.es

Fecha de recepción: septiembre de 2022

Fecha de aceptación: enero de 2023

Resumen

El cambio de polaridad se considera uno de los problemas más desafiantes dentro del ámbito de la minería de opiniones. Los operadores de cambio de polaridad son elementos lingüísticos contextuales que pueden aumentar, reducir o neutralizar la polaridad de una palabra denominada “foco” que se incluye en la expresión de una opinión. La detección automática de tales elementos mejora el rendimiento y la precisión de los sistemas computacionales para la minería de opiniones. Desde un enfoque simbólico, este artículo presenta un avance en el procesamiento automático de los operadores contextuales que afectan la polaridad de las opiniones expresadas en tuits escritos en inglés. Con este fin, proponemos un nuevo modelo basado en el conocimiento relativo a la cuantificación en inglés, la cual puede incrementar o reducir la polaridad de las opiniones. En concreto, describimos las reglas lingüísticas de cada categoría de operador asociado a la cuantificación, incluyendo información sobre el alcance y la dirección del operador con respecto al foco. Además, indicamos las fórmulas matemáticas que calculan la fuerza del efecto sobre la polaridad previa. Finalmente, presentamos las matrices asociadas a las reglas lingüísticas, las cuales facilitan la modelización del conocimiento en un sistema de minería textual.

Palabras clave: minería de opiniones; análisis del sentimiento; cambio de polaridad; intensificación; cuantificación.

Abstract

Polarity shifting can be considered one of the most challenging problems in the context of sentiment analysis. Polarity shifters are treated as linguistic contextual items that can increment, reduce or neutralise the polarity of a word called ‘focus’ included in an opinion. The automatic detection of such items enhances performance and accuracy of computational systems for opinion mining. From a symbolic approach, we aim to advance in the automatic processing of the polarity shifters that affect the opinions expressed on tweets in English. To this end, we describe a novel knowledge-based model to deal with quantification in English, which increments or reduces the polarity of opinions. In particular, we explain the linguistic rules of each category of quantification shifter, including information about the scope and direction with respect to the focus. Furthermore, we present the mathematical formulae that calculate the strength of the effect on the prior polarity. Finally, we describe the matrices associated to the linguistic rules, which serve to model the knowledge in text-mining systems.

Keywords: opinion mining; sentiment analysis; polarity shifting; intensification; quantification.

1. Introducción

El desarrollo de la Web 2.0 ha supuesto no solo un cambio en la forma de comunicarnos dentro de un nuevo marco discursivo (Álvarez, 2011; Mancera y Pano, 2013; Escandell Vidal, 2014), sino también el surgimiento de múltiples plataformas, como las redes sociales (e.g. Twitter), que sirven de altavoz a las opiniones generadas por sus usuarios. De esta manera, la detección automática de opiniones escritas en las redes sociales resulta hoy crucial para aquellas empresas que quieren mejorar no solo la calidad de sus productos o servicios, sino también la atención a sus clientes y su competitividad. Por esta razón, la minería de opiniones se ha convertido en una disciplina de investigación de enorme interés en la industria y en la academia por sus múltiples aplicaciones sociales y empresariales¹. No obstante, aunque se ha producido un gran avance en la última década, el grado de precisión y optimización de los sistemas computacionales basados en la minería de opiniones requiere en general de notables mejoras, entre las que se encuentran la detección y el tratamiento automáticos de los operadores contextuales que cambian la polaridad inicial (positiva o negativa) de los comentarios de los usuarios.

En este contexto, el objetivo de este artículo es describir la construcción de una arquitectura modular desde un enfoque simbólico que facilite el procesamiento computacional de los diversos tipos de operadores contextuales incluidos en la dimensión de la cuantificación en inglés. El estudio sobre la detección automática de estos operadores se realizó a partir del análisis del conocimiento lingüístico teórico, el cual se contrastó posteriormente con evidencias lingüísticas provenientes de los mensajes publicados en Twitter. En cuanto a la metodología empleada en este estudio, cabe señalar que los dos enfoques de investigación que suelen adoptarse en la minería de opiniones y, por extensión, en el tratamiento de los operadores contextuales son el estadístico (Moreno-Ortiz y Fernández-Cruz, 2015; Cruz y otros, 2016; Jiménez-Zafra y otros, 2020; Schulder y otros, 2020) y el simbólico (Taboada y otros, 2011; Perriñán-Pascual y Arcas-Túnez, 2019).

El primer enfoque, basado en corpus y el más habitual en los estudios de minería textual, se caracteriza por la construcción de sistemas en los que se aplican técnicas matemáticas de aprendizaje automático. Estas técnicas se fundamentan principalmente en modelos supervisados, con lo cual requieren de un entrenamiento previo a través de corpus etiquetados con el fin de detectar patrones e inferir así conocimiento de forma automática. En cambio, el segundo enfoque está basado directamente en el conocimiento lingüístico, es decir, recurre al almacenamiento y la reutilización de recursos léxicos y a la construcción de reglas lingüísticas.

1 Una breve introducción sobre la minería textual, y en particular sobre la minería de opiniones como una de sus áreas de investigación, puede encontrarse en Perriñán-Pascual (2022).

En este sentido, nuestro objetivo final es elaborar una base de conocimiento lingüístico suficientemente robusta que permita mejorar la actuación de los sistemas computacionales basados en el procesamiento del lenguaje natural sin necesidad de crear costosos recursos adicionales, como los extensos corpus que se utilizan en el entrenamiento de los sistemas de clasificación textual.

Asimismo, los objetivos específicos que pretendemos con este trabajo son, por una parte, profundizar en los operadores contextuales de cambio de polaridad incluidos dentro del fenómeno de la cuantificación en inglés, con el propósito de que cada categoría pueda recibir un tratamiento específico en la arquitectura al disponer de una información detallada sobre la posición que puede ocupar con respecto a su foco (i.e. dirección y alcance), el tipo de variación en la polaridad inicial que puede generar y su cuantificación. Por otra parte, también pretendemos elaborar un modelo replicable para el procesamiento de otras lenguas distintas al inglés, con un elevado grado de flexibilidad y trazabilidad, constituyendo, de esta manera, un sistema relativamente sencillo de mantener y optimizar, bien sea añadiendo nuevas reglas para aumentar su cobertura, bien sea corrigiendo o eliminando reglas existentes para mejorar la precisión.

2. El cambio de polaridad en la minería de opiniones: la cuantificación en inglés

Tradicionalmente, la metodología empleada para detectar la polaridad en la minería de opiniones se basa en los lexicones de polaridad, i.e. listas de palabras anotadas con respecto al sentimiento expresado (i.e. positivas, negativas o neutras). En aquellos casos en que la clasificación de la opinión se realiza únicamente en base a estas palabras, los resultados son poco satisfactorios, puesto que se tiene en cuenta solamente el significado léxico de estos términos y no el contexto en el que aparecen. Por esta razón, Polanyi y Zaenen (2004) consideraron esta metodología como incompleta, proponiendo que las unidades léxicas con significado actitudinal pudieran experimentar un cambio en su valencia (i.e. polaridad). De esta forma, una palabra puede pasar de tener una “valencia simple” a una “valencia determinada por el contexto”, la cual se genera por la aparición en el texto de determinados operadores contextuales de cambio de polaridad (en inglés, *contextual valence shifters*).

A partir de este estudio, fueron apareciendo otros trabajos que propusieron diversos modelos en el tratamiento computacional de los operadores contextuales aplicados a la minería de opiniones. Una gran mayoría de estos estudios se centran en la negación en inglés, pero son cada vez más los autores que señalan la conveniencia de incorporar más tipos de operadores en los modelos computacionales (Li y otros, 2010; Li y otros, 2013). En este sentido, la cuantificación constituye un tipo de operador contextual que, a pesar de ser menos estudiado en comparación con la negación, también modifica la polaridad inicial de una opinión, incrementando (i.e. intensificadores) o disminuyendo (i.e. atenuadores) su

intensidad (Quirk y otros, 1985; Rhee, 2016). Por ejemplo, en la opinión *the room service at the hotel was very good*, la polaridad de *good* junto con *very* es más intensa (i.e. valencia determinada por el contexto) que la polaridad de *good* solo (i.e. valencia simple).

Entre los trabajos relacionados con la cuantificación en inglés, destacamos en primer lugar los modelos de Polanyi y Zaenen (2004) y de Kennedy e Inkpen (2006), que calcularon el cambio de intensidad sumando o restando un número fijo a la valencia inicial del término, en lugar de invertir la polaridad como contemplaban para la negación. Por una parte, en Polanyi y Zaenen (2004: 108) se suma un punto en caso de ser un intensificador el que precede al término de sentimiento y se resta un punto en caso de ser un atenuador el que lo precede. Por otra parte, Kennedy e Inkpen (2006) implementaron las mismas propuestas de Polanyi y Zaenen (2004). En concreto, asignaron un valor por defecto de +2 a todo término con polaridad positiva y, si van precedidos de un intensificador o atenuador, se les da una polaridad final de +3 o +1, respectivamente; igualmente, asignan un valor de -2 a todo término con polaridad negativa y, si van precedidos de un intensificador o atenuador, se les da una polaridad final de -3 o -1, respectivamente. No obstante, no consiguieron solucionar los casos en los que se combina intensificación con negación, como en *not very good*, porque el sistema actúa de la siguiente manera: *not* (-2) combinado con *very* genera un valor de -3 que multiplicado por la polaridad de *good* (+2) es igual a -6, siendo un valor negativo muy elevado y, por tanto, erróneo.

Kennedy e Inkpen (2006) modificaron su modelo inicial y crearon atributos artificiales para incorporar estos operadores en los sistemas de clasificación de aprendizaje automático. Por ejemplo, crearon *int_good* para el intensificador de *good*. Wilson y otros (2009) mejoraron este modelo, añadiendo atributos como *preceded_by_adverb* (i.e. si la palabra precedente es un adverbio distinto de *not*), *preceded_by_intensifier* (i.e. si la palabra precedente es un intensificador) y *self_intensifier* (i.e. si la palabra misma es un intensificador), entre otros muchos. Morsy y Rafea (2012), basándose en este modelo de Kennedy e Inkpen (2006), optimizaron el inventario de atributos, reduciendo su número considerablemente.

Por otro lado, Whitelaw y otros (2005) propusieron un nuevo método de clasificación de sentimiento basado en la extracción y el análisis de lo que denominan “grupos de valoración”, que se representan por un conjunto de atributos dentro de una taxonomía semántica basada en la Teoría de la Valoración de Martin y White (2005). Un grupo de valoración está compuesto por un “adjetivo nuclear” (e.g. *happy*), que representa la actitud tipo que define al grupo, con una lista de modificadores de valoración que lo pueden preceder, cada uno de los cuales denota una transformación de uno o más de los atributos de ese término nuclear. De esta forma, los intensificadores se consideran modificadores del atributo “intensidad”, dentro, a su vez, del atributo “gradación”, para tres niveles de intensidad contemplados: neutra, alta y leve. Por ejemplo, si a *happy* se le asigna un nivel de intensidad neutra, con el intensificador *very* se transforma en una intensidad alta y, si lo combinamos

con la negación *not* (i.e. *not very happy*), pasa a una intensidad leve. Carrillo-de-Albornoz y Plaza (2013: 1620) consideraron que una limitación de este modelo es que asignan, de forma errónea, un valor fijo a todos los intensificadores y atenuadores “[...] but it is clear that not all intensifiers amplify or diminish polarity/emotions to the same extent”. Para superar esta limitación, surgieron otros modelos basados en listas de intensificadores con porcentajes asociados representativos del efecto que producen, positivos en el caso de los intensificadores y negativos en el caso de los atenuadores (Taboada y otros, 2011; Phu y Tuoi, 2014; Taboada, 2016). Por ejemplo, en Taboada y otros (2011) y Taboada (2016) se propusieron valores como +100 % para *(the) most*, +50 % para *extraordinary*, +25 % para *really*, +15 % para *very*, -10 % para *pretty*, -30 % para *somewhat* y -50 % para *slightly*. Sin embargo, no se detallaron los criterios que se siguieron en la asignación de estos porcentajes.

Otros modelos publicados son los basados en reglas composicionales. Por ejemplo, Liu y Seneff (2009) presentaron una visión unificada de los operadores contextuales (i.e. intensificadores, atenuadores y negación) basada en una serie de reglas composicionales. Neviarouskaya y otros (2010, 2015) construyeron un modelo basado en un lexicón y una serie de reglas composicionales. En el lexicón incluyeron tres grupos de elementos: los términos actitudinales, los modificadores y las palabras funcionales y, finalmente, los operadores modales. Así, dentro de las palabras funcionales, podemos encontrar adjetivos (e.g. *rising*), nombres (e.g. *increase*) y verbos (e.g. *grow*) que incrementan la intensidad de las palabras actitudinales asociadas. Además, dentro de los modificadores, los autores incluyeron intensificadores como los adverbios de grado (e.g. *significantly* y *slightly*) y los adverbios de afirmación (e.g. *absolutely* y *seemingly*), los cuales tienen un impacto directo sobre la intensidad del rasgo actitudinal de las palabras. Los anotadores asignan a estos adverbios coeficientes representativos de su grado de intensidad (i.e. intensificación o atenuación), dentro de un rango de 0,0 a 2,0 (e.g. *slightly* tiene asignado un coeficiente de 0,2, y *perfectly*, de 1,9). Igualmente, establecieron distintas reglas composicionales, dentro de las cuales incluyeron la denominada “regla de intensificación”, que implica aumentar o disminuir el valor numérico de la polaridad, aplicándose en dos casos específicos: si un adverbio de grado o de afirmación se asocia a un término actitudinal (e.g. $Pos_score(\textit{extremely happy}) > Pos_score(\textit{happy})$), o si un adjetivo o adverbio se usa en su grado comparativo o superlativo (e.g. $Neg_score(\textit{sad}) < Neg_score(\textit{sadder}) < Neg_score(\textit{saddest})$).

Siguiendo esta línea, Yoo y Nam (2018), dentro de su modelo híbrido de reglas composicionales para el chino, contemplaron para el tratamiento de estos operadores la regla de añadir +1 al valor numérico de la polaridad en caso de tratarse de un intensificador asociado a una palabra actitudinal y de restar -1 en caso de tratarse de un atenuador.

Por su parte, Asghar y otros (2017) establecieron una fórmula para calcular el valor de la polaridad modificada por el efecto de los intensificadores. Con este fin, partieron de una lista de intensificadores a los que asignaron un valor numérico positivo o negativo (e.g. asigna-

ron un valor de +0,5 al intensificador *very*, positivo por intensificar la polaridad), y, teniendo en cuenta la polaridad inicial de la palabra de sentimiento, aplicaban la fórmula por la que se multiplica el valor asignado al intensificador por la polaridad inicial y luego se vuelve a sumar la polaridad inicial. Por ejemplo, si *comfortable* tiene asignada una polaridad inicial de +0,45 y el intensificador *very* tiene un valor de +0,5, el bigrama *very comfortable* tendría una polaridad final de 0,675, la cual resulta de $((0,45 \text{ [polaridad inicial]} * 0,5 \text{ [valor de very]}) + 0,45 \text{ [polaridad inicial]})$.

Finalmente, caben mencionar los estudios de Liu (2015) y Varathan y otros (2017), basados en reglas de comparación dentro de la minería de opiniones, ya que esta podría considerarse una línea de investigación adicional dentro del ámbito de la cuantificación (e.g. una opinión como *Coke tastes the best among all soft drinks* sería equivalente a *Coke tastes very good*).

Por otro lado, Carrillo-de-Albornoz y Plaza (2013) y Sintsova y otros (2018) contemplaron los intensificadores como modificadores de las emociones, llegando a la conclusión de que el efecto sobre ellas por parte de los intensificadores es menor que en el caso de la negación o la modalidad. Sintsova y otros (2018) recopilaron una lista de intensificadores (e.g. *much*) y otra de atenuadores (e.g. *a bit*) y, a partir de estos, añadieron los n-gramas que los contienen. Posteriormente, clasificaron cada modificador según su posición con respecto al término con rasgos emocionales: delante (e.g. *lots of*), detrás (e.g. *very much*) o ambos (e.g. *less*). El alcance se fijó en una palabra delante o detrás, según la posición definida para cada uno de ellos.

Otro modelo que merece la pena destacar es el presentado por Chen y Zhu (2014), el cual aplicaron a la traducción automática. En él, establecieron un esquema de cómo se calcula el valor numérico de la polaridad modificada, combinando tres tipos de términos almacenados en lexicones: las palabras de sentimiento, los intensificadores y las palabras de negación. A cada uno de estos tipos de términos le asignaron unos valores numéricos en función de su intensidad. De esta forma, las palabras de sentimiento pueden tener los valores -0,8, -0,4, 0, 0,4 y 0,8; los intensificadores, -0,5, 0,5, 0,7 y 0,9, y las palabras de negación, -0,8 y -0,6. Para calcular la polaridad que resulta por el efecto de los operadores, se utilizó la fórmula $(1 + (\text{valor negación}) * (\text{valor intensificador})) * \text{valor palabra sentimiento}$. Por ejemplo, para *not very good*, siendo la polaridad inicial de cada palabra de -0,8, +0,7 y +0,4, respectivamente, el valor de la polaridad final sería 0,176, i.e. $(1 + (-0,8) * (0,7)) * 0,4$. Para detectar automáticamente el modificador a partir de la palabra de sentimiento, establecen una ventana fija de 7 palabras.

Finalmente, concluimos este apartado con otros estudios que han contribuido al tratamiento computacional de la cuantificación en inglés en el marco de los sistemas de minería textual. Musat y Trausan-Matu (2010), quienes estudiaron la combinación de palabras de

negación y atenuadores en el ámbito económico, concluyeron que los atenuadores no solo modifican la polaridad de la palabra de sentimiento asociada, sino que la invierten completamente. Novak y otros (2015: 13) consideraron el papel de los emojis como amplificadores y modificadores de significado. Ruppenhofer y otros (2015) realizaron una clasificación jerárquica automática de los adverbios en inglés según su efecto intensificador o atenuador utilizando metadatos de *ratings* sobre las valoraciones de productos. Chen y otros (2019) incorporaron atributos de negación y de expresiones de intensidad (tan solo *very* y *so*) en un modelo de redes neuronales de aprendizaje profundo con el objeto de intentar solucionar el problema de la dependencia de dominios y de lexicones externos propio de los métodos estadísticos. Mudgal y Khunteta (2020) publicaron un estudio general sobre el tratamiento de la doble intensificación, principalmente en contextos negativos, con un modelo de reglas básicas y atributos incorporado dentro de un sistema de clasificación.

3. Modelo propuesto para el procesamiento automático de la cuantificación

Nuestro modelo se basa en la construcción de una arquitectura modular que permite la detección y el tratamiento computacional de los tipos de operadores contextuales incluidos en la dimensión específica de la cuantificación. A continuación, describimos dos aspectos fundamentales de nuestro modelo (i.e. la escala de grados de polaridad y la tipología de operadores contextuales), además de los principales componentes que configuran nuestra arquitectura modular.

3.1. Escala de grados de polaridad en el lexicon de sentimiento

Esta arquitectura requiere una escala previa de grados de polaridad en un lexicon de sentimiento incluido en el sistema computacional donde se implemente. En este sentido, consideramos que, en primer lugar, el sistema computacional debe identificar la palabra o conjunto de palabras que están incluidas dentro de un lexicon de sentimiento, construido a partir de recursos ya existentes como SentiWordNet (Esuli y Sebastiani, 2006). En este tipo de recursos, estas palabras dotadas de sentimiento llevan asociado un tipo de polaridad inicial (positiva o negativa), así como un grado de intensidad de esa polaridad. Este grado de intensidad viene identificado por un número dentro de una escala lineal predefinida, acompañado de un signo + o - según se trate de una polaridad positiva o negativa, respectivamente. Independientemente del recurso utilizado, esta escala numérica predefinida del lexicon es traducida de forma automática a otra de tres grados de intensidad (i.e. muy fuerte, fuerte y leve) para cada uno de los polos o sentidos de polaridad (i.e. positivo o negativo), es decir, un rango de [-3, -2, -1, 0, +1, +2, +3]. Los tres grados especifican la intensidad de la ausencia o presencia de la cualidad, valoradas en relación con la magnitud que el hablante presupone. En la tabla 1, se muestra esta escala de polaridad con las distintas graduaciones de intensidad.

TABLA 1

Escala de grados de polaridad del lexicón de sentimiento

| ESCALA | GRADO | TIPO POLARIDAD E INTENSIDAD | EJEMPLOS |
|------------------|-------|--|---------------------------|
| SENTIDO NEGATIVO | -3 | Polaridad negativa con intensidad muy fuerte | <i>detest, horrible</i> |
| | -2 | Polaridad negativa con intensidad fuerte | <i>dislike, bad</i> |
| | -1 | Polaridad negativa con intensidad leve | <i>displease, so-so</i> |
| | 0 | Ausencia de polaridad | |
| SENTIDO POSITIVO | +1 | Polaridad positiva con intensidad leve | <i>please, acceptable</i> |
| | +2 | Polaridad positiva con intensidad fuerte | <i>like, good</i> |
| | +3 | Polaridad positiva con intensidad muy fuerte | <i>love, excellent</i> |

A este grado de polaridad e intensidad inicial lo denominamos Polaridad de Entrada.

3.2. Tipología de operadores contextuales de cambio de polaridad asociados a la cuantificación

Los Modificadores de Polaridad son operadores que pueden modificar, en sentido positivo o negativo, la Polaridad de Entrada que ofrece el sistema en un primer momento. A la polaridad resultante del efecto de estos modificadores la hemos denominado Polaridad Modificada. Estos modificadores pueden clasificarse a su vez en dos tipos:

- a) Los Modificadores Intensificadores aumentan el grado de una determinada polaridad (en sentido positivo o negativo). Para que un operador pueda considerarse un intensificador, debe tenerse en cuenta el efecto final que produce en la polaridad, el cual puede ser un incremento de intensidad o graduación en el mismo sentido al de la Polaridad de Entrada, o un incremento en sentido opuesto.
- b) Los Modificadores Atenuadores disminuyen una determinada polaridad (en sentido positivo o negativo). Igualmente, para que un operador pueda considerarse un atenuador, debe tenerse en cuenta el efecto final que produce, que en este caso consiste en una reducción de intensidad o graduación en la escala de polaridad en el mismo sentido al de la Polaridad de Entrada.

Finalmente, para medir la intensidad en este aumento o reducción generados por los Modificadores Intensificadores y los Modificadores Atenuadores, se han contemplado dos escalas de intensidad de la variación que facilitan la cuantificación de este efecto: intensidad leve e intensidad fuerte. En la tabla 2, exponemos brevemente el esquema de funcionamiento de estos dos tipos de operadores contextuales.

TABLA 2

Tipología de operadores contextuales de cuantificación según tipo, variación e intensidad de polaridad

| TIPOLOGÍA SEGÚN TIPO DE VARIACIÓN | VARIACIÓN | INTENSIDAD VARIACIÓN | EFFECTO FINAL EN POLARIDAD | POLARIDAD FINAL |
|-----------------------------------|--|--|--|----------------------|
| Modificadores Intensificadores | Efecto final: aumentan Polaridad Entrada (positiva o negativa) | <ul style="list-style-type: none"> • Leve • Fuerte | <ul style="list-style-type: none"> • Positiva $\Rightarrow \Delta$ Positiva • Negativa $\Rightarrow \Delta$ Negativa • Positiva $\Rightarrow \Delta$ Negativa • Negativa $\Rightarrow \Delta$ Positiva | Polaridad Modificada |
| Modificadores Atenuadores | Efecto final: reducen Polaridad Entrada (positiva o negativa) | <ul style="list-style-type: none"> • Leve • Fuerte | <ul style="list-style-type: none"> • Positiva $\Rightarrow \nabla$ Positiva • Negativa $\Rightarrow \nabla$ Negativa | Polaridad Modificada |

3.3. Componentes principales de la arquitectura modular

Los componentes básicos de la arquitectura modular de los operadores contextuales que hemos estudiado son la dimensión, las categorías de los operadores contextuales, las reglas lingüísticas, las matrices de las reglas lingüísticas y las fórmulas de variación.

3.3.1. Dimensión

La dimensión constituye el módulo donde se incluyen los operadores contextuales de cambio de polaridad asociados a un fenómeno lingüístico. Para el objeto de estudio de este artículo, la dimensión es CUANTIFICACIÓN. A su vez, cada dimensión de operadores incluye una serie de categorías de operadores, reglas lingüísticas, matrices y fórmulas.

3.3.2. Categorías de operadores contextuales

Dentro de la dimensión CUANTIFICACIÓN, los operadores contextuales se clasifican en categorías según los patrones similares de funcionamiento que presentan y el efecto que producen en la Polaridad de Entrada. Debe recordarse que no todos los operadores de cuantificación amplifican o atenúan la polaridad con la misma intensidad.

3.3.3. Reglas lingüísticas

Cada una de las categorías de operadores definidas en el punto anterior tiene asociada una regla lingüística que explica la variación que se produce en la polaridad. Ya que estas reglas

expresan la relación causa-efecto que se produce entre los operadores de la categoría y la variación de la polaridad, estas adoptan la forma de *Si...*, *entonces...* Para construir las reglas lingüísticas se tienen en cuenta los distintos atributos de los operadores:

1. El foco del operador, i.e. el elemento que se ve afectado directamente por el operador. Es la palabra que se incluye en el lexicón de sentimiento y la que determina, sin el efecto de los operadores contextuales, la polaridad inicial del enunciado. En cada regla lingüística se especifica uno o varios focos, cada uno de los cuales viene definido por los siguientes atributos:
 - a) La categoría gramatical del foco, que está asociada a un adjetivo, verbo o adverbio para el caso de la cuantificación.
 - b) La Polaridad de Entrada, cuyo valor se encuentra en el rango [-3, +3].
 - c) El efecto que el operador ocasiona en la Polaridad de Entrada, el cual se expresa en términos de si cambia el sentido de su polaridad inicial y con qué intensidad (i.e. leve o fuerte).
2. El alcance del operador, i.e. el número máximo de palabras que puede haber desde el foco hasta el operador contextual.
3. La dirección del operador, i.e. la dirección que debe seguirse para medir el alcance del operador. Es decir, indica si el operador está antepuesto al foco (dirección izquierda) o si está pospuesto al foco (dirección derecha).

3.3.4. Matrices de las reglas lingüísticas

Cada regla lingüística se representa en una matriz con el fin de facilitar su incorporación al sistema como regla computacional. En la tabla 3, tabla 4 y tabla 5, se muestran los cinco bloques que componen cada matriz:

TABLA 3

Bloques I y II del modelo de matriz de reglas lingüísticas

| ELEMENTOS (1/3) | ETIQUETA | EXTENSIÓN ETIQUETA | DESCRIPCIÓN |
|--|------------------------|--------------------|---|
| I. DIMENSIÓN E INICIO DE ENUNCIADO | <CUANT_Enunciado> | | Indica el inicio de un enunciado de dimensión CUANTIFICACIÓN. Lleva asociada una etiqueta de cierre. |
| II. CATEGORÍA de OPERADORES CONTEXTUALES | ENG_CUANT_Xn n=1,...,n | "OPERADOR (Y/O)" | Identificación de categoría de operadores dentro de la dimensión CUANTIFICACIÓN para el inglés. Cada categoría incluye un conjunto de operadores que tienen efectos sobre el foco y funcionamientos similares. Los operadores: <ul style="list-style-type: none"> • Pueden ser una palabra, un grupo de palabras. Pueden hacer referencia a una o más posibilidades con los operadores booleanos Y (han de darse todos los elementos) y O (ha de darse al menos uno de ellos). • Normalmente, van entre comillas [" "], sin tener en cuenta minúsculas ni mayúsculas. • La etiqueta de la categoría, que la identifica, está formada por tres parámetros: la lengua (ENG), la dimensión (CUANT) y su variable de identificación (X_n) que llevará un subíndice para cada categoría dentro de la dimensión. |
| Atributos OPERADOR | ATRIBUTOS OPERADOR | | Define y delimita el conjunto de atributos de la categoría de operadores contextuales identificada. |
| | ALCANCE | NÚMERO | Indica el número de palabras o posiciones máximas que puede haber desde el operador contextual hasta el foco (palabra dentro de lexicón de sentimiento) que resulta afectado por él. Es un número. |
| | DIRECCIÓN | DCHA/IZQ | Dirección del alcance. Dirección que se sigue al contar el número de palabras desde foco hasta el operador contextual. Puede ser hacia la derecha (DCHA) o hacia la izquierda (IZQ). Si el operador va antepuesto al foco, la dirección será IZQ; si va pospuesto al foco, la dirección será DCHA. |

TABLA 4

Bloque III del modelo de matriz de reglas lingüísticas

| ELEMENTOS (2/3) | ETIQUETA | EXTENSIÓN ETIQUETA | | EXTENSIÓN ETIQUETA | DESCRIPCIÓN |
|---|---------------------------|----------------------------|-----|----------------------------|--|
| III. INICIO CONDICIONES DE LA REGLA LINGÜÍSTICA | SI | CASO 1 | Y/O | CASO 2 | Delimita el inicio del bloque de condiciones de la regla lingüística, que están relacionadas con los atributos del foco del operador contextual. Puede haber una condición o varias, que aparecen unidas por los booleanos Y u O. Cada una de las condiciones vendrá identificada por la extensión CASO y un número correlativo. |
| III. 1. FOCO del OPERADOR CONTEXTUAL | ENG_CUANT_Fn n=1,...,n | | | | Identificación del foco asociado a una categoría de operadores de la dimensión CUANTIFICACIÓN en inglés, especificada en los bloques I y II de la matriz: <ul style="list-style-type: none"> El foco siempre vendrá definido por una categoría gramatical (CATG), con una polaridad e intensidad definida de entrada. El foco es el elemento objeto de la variación de polaridad por el efecto del operador. Al igual que ocurre con la categoría de operadores, su etiqueta está formada por tres parámetros: la lengua (ENG), la dimensión (CUANT) y su variable de identificación (F_n) que llevará un subíndice n para cada foco. Este subíndice coincide con el número asignado a la categoría de operadores que hacen variar ese foco. Por ejemplo, ENG_CUANT_F1 irá asociado al operador ENG_CUANT_X1. |
| Atributos del FOCO | ATRIBUTOS FOCO | | | | Delimita el bloque de atributos del foco identificado para cada categoría. |
| | CATG | N/JJ/VB/RB | Y/O | N/JJ/VB/RB | Indica la categoría gramatical del foco (incluido en el lexicon sentimiento). Puede haber más de una posibilidad y cada una de ellas se asocia a un caso (N=NOMBRE/ JJ=ADJETIVO/ VB=VERBO/ RB=ADVERBIO). |
| | POLARIDAD_E | [POSITIVA/NEGATIVA, GRADO] | | [POSITIVA/NEGATIVA, GRADO] | Polaridad de Entrada asociada al foco, sin el efecto del operador contextual. Viene identificada por su sentido (positiva o negativa), y por su graduación dentro del rango [-3, +3]. |

TABLA 5

Bloques IV y V del modelo de matriz de reglas lingüísticas

| ELEMENTOS (3/3) | ETIQUETA | EXTENSIÓN ETIQUETA | DESCRIPCIÓN |
|---|-------------------------|---------------------------------|--|
| IV. EFECTOS DEL CUMPLIMIENTO DE LA CONDICIÓN (CASO) | ENTONCES | EFECTO SOBRE POLARIDAD_E (FOCO) | Delimita el bloque de los efectos del operador contextual sobre la Polaridad de Entrada del Foco. Cada efecto irá asociado a cada condición o caso (CASO) detallado en el bloque III anterior. |
| IV. 1. EFECTO MODIFICADOR | POLARIDAD_MODIFICADOR | SÍ/NO | Identifica si el operador actúa como Modificador (SÍ) o no (NO). |
| | T_POLARIDAD_MODIFICADOR | INTENSIFICADOR/ATENUADOR | Si actúa como Modificador (SÍ), clasifica el tipo: INTENSIFICADOR o ATENUADOR. |
| | CAMBIA_POLARIDAD | SÍ/NO | Señala si cambia el sentido de la Polaridad de Entrada (POLARIDAD_E) del foco (SÍ) o no (NO). |
| | INTENSIDAD_NIVEL | LEVE/FUERTE | Identifica el grado de intensidad de la variación efectuada por el Modificador. Dos niveles posibles: Leve (LEVE) o Fuerte (FUERTE). |
| | FÓRMULA | POLARIDAD (1) | Identifica la fórmula que debe aplicarse. |
| | POLARIDAD_M | [POSITIVA/NEGATIVA, GRADO] | Identifica la Polaridad Modificada o de salida por el efecto del modificador. Viene identificada por su sentido (positiva o negativa), y por su graduación dentro del rango [-3, +3], al igual que la Polaridad de Entrada (Polaridad_E) del foco. |
| V. DIMENSIÓN Y FINAL DE ENUNCIADO | | <CUANT_Enunciado/> | Indica el final de un enunciado de dimensión CUANTIFICACIÓN. Lleva asociada una etiqueta de inicio. |

3.3.5. Fórmulas de variación

Cada matriz de regla lingüística lleva asociada una o varias fórmulas, cuyo fin es realizar el cálculo de la Polaridad Modificada teniendo en cuenta las siguientes variables:

- La Polaridad de Entrada de la palabra, que viene determinada por el sentido de la polaridad y el grado de intensidad.
- Si la variación se produce en el mismo sentido al de la Polaridad de Entrada (i.e. de positiva a positiva o de negativa a negativa), o en sentido opuesto (i.e. de positiva a negativa o de negativa a positiva).

- Si se produce un aumento o una reducción del sentimiento.
- La intensidad de la variación que se produce (i.e. leve o fuerte).

En la figura 1, se presentan de forma gráfica todos los componentes de la arquitectura, así como sus relaciones.

FIGURA 1

Componentes básicos arquitectura operadores contextuales CUANTIFICACIÓN



4. Metodología

4.1. Operadores contextuales de CUANTIFICACIÓN

En primer lugar, partimos del concepto de opinión como enunciado lingüístico, centrándonos en la exploración de contenidos textuales generados por los usuarios de redes sociales (e.g. Twitter). A partir del conocimiento lingüístico teórico para el inglés (Quirk y otros, 1985; Mackenzie y Martínez-Caro, 2012), consideramos que, en la expresión de la opinión, la cuantificación viene representada por medio de adverbios de cantidad o de grado que pueden modificar tres posibles categorías de palabras incluidas en el lexicon de sentimiento del sistema computacional: verbo, adverbio y adjetivo. A continuación, modelizamos las posibles estructuras sintácticas en las que estos adverbios modificadores pueden aparecer en un enunciado de opinión con el fin de descubrir la posición

que estos ocupan con respecto al foco, además de su dirección y alcance. Tras explorar estos patrones sintácticos en tuits, concluimos que podemos encontrar un operador de la dimensión CUANTIFICACIÓN hasta en un máximo de cinco posiciones a la derecha cuando el foco es un verbo y de una posición a la izquierda cuando es un adverbio o adjetivo. Igualmente, se evidenció que estos operadores pueden modificar la polaridad inicial del foco, aumentándola o atenuándola. El siguiente paso consistió en realizar una selección de los operadores que el sistema debería detectar de forma automática. De esta selección, se obtuvieron cuatro categorías de operadores.

Por una parte, la categoría ENG_CUANT_X1 incluye operadores como *very*, *really*, *quite* y *extremely*, entre otros, y la categoría ENG_CUANT_X2 contiene operadores como *very much* y *a lot*. Mientras los operadores de la primera categoría siempre aparecen antepuestos al foco (adjetivo o adverbio) en un rango de 1 posición, los operadores de la segunda categoría siempre aparecen pospuestos al foco (verbo) dentro un alcance máximo de 5 posiciones. En ambas categorías, la Polaridad de Entrada del foco puede ser positiva o negativa. Si la Polaridad de Entrada es positiva, el grado de intensidad puede ser leve (+1) o fuerte (+2), donde la Polaridad Modificada será también positiva, con intensidad fuerte (+2) en el caso de intensidad inicial leve, y muy fuerte (+3) para el caso de intensidad inicial fuerte. Del mismo modo, si la Polaridad de Entrada es negativa, el grado de intensidad también puede ser leve (-1) o fuerte (-2), donde la Polaridad Modificada será también negativa, con intensidad fuerte (-2) en el primer caso, y muy fuerte (-3) en el segundo caso. Por tanto, los operadores de ambas categorías son Modificadores Intensificadores que producen una variación en el mismo sentido con una intensidad leve.

Por otra parte, la categoría ENG_CUANT_X3 incluye operadores como *a bit*, *a little*, *just a bit* y *slightly*, entre otros, mientras que la categoría ENG_CUANT_X4 incorpora operadores como *very little*. Los operadores de ambas categorías pueden ir antepuestos o pospuestos al foco, dependiendo de si se trata de un adjetivo o un verbo, respectivamente. Cuando el foco es un verbo, el alcance máximo será de 5 posiciones, pero, cuando el foco es un adjetivo, el rango será de 1 posición. En la categoría ENG_CUANT_X3, la Polaridad de Entrada del foco puede ser positiva o negativa. Si es positiva, el grado de intensidad puede ser leve (+1) o fuerte (+2), donde la Polaridad Modificada será negativa con intensidad leve (-1) en el caso de intensidad inicial leve, y positiva con intensidad leve (+1) para el caso de intensidad inicial fuerte. De igual manera, si la Polaridad de Entrada es negativa, el grado de intensidad puede ser fuerte (-2), donde la Polaridad Modificada será también negativa con intensidad leve (-1). Por tanto, los operadores de esta categoría son Modificadores Intensificadores, si la Polaridad de Entrada es positiva con intensidad leve (+1), que producen una variación en sentido opuesto al de esa Polaridad de Entrada con una intensidad leve, y Modificadores Atenuadores, en el caso de Polaridad de Entrada con intensidad fuerte, tanto positiva (+2) como negativa (-2), que producen una variación en el mismo sentido al de esa Polaridad de Entrada con una intensidad leve. En la categoría ENG_CUANT_X4, la Polaridad de Entrada del

foco puede ser positiva, con intensidad leve (+1) o fuerte (+2), donde la Polaridad Modificada será negativa, pero con un grado de intensidad leve (-1) en el caso de que el grado de la Polaridad de Entrada sea leve y con un grado de intensidad fuerte (-2) en el caso de que el grado de la Polaridad de Entrada sea fuerte. Por tanto, los operadores de esta categoría son Modificadores Intensificadores que producen una variación en sentido opuesto al de la Polaridad de Entrada con una intensidad leve o fuerte.

A continuación, presentamos ejemplos que ilustran las cuatro categorías anteriores de operadores, donde se aplican las correspondientes fórmulas para el cálculo de la Polaridad Modificada. PE y PM_CUANT representan la Polaridad de Entrada y la Polaridad Modificada, respectivamente:

- (1) Categoría ENG_CUANT_X1
Denny's Restaurant on #Yelp: The food was good, the service was quite acceptable, the only downside was the bathroom
acceptable [PE: +1], *quite acceptable* [PM_CUANT: +1+1= +2].
- (2) Categoría ENG_CUANT_X2
It displeases me very much when professors don't email me back
displease [PE: -1], *displease...very much* [PM_CUANT: -1-1= -2].
- (3) Categoría ENG_CUANT_X3
i know it's a fake, but i like it a bit.
like [PE: +2], *like...a bit* [PM_CUANT: +2-1=+1].
- (4) Categoría ENG_CUANT_X4
There is very little amusing about living in Nuneaton.
amusing [PE: +2], *very little amusing* [PM_CUANT: (+2 * 0)-2= 0-2=-2].

4.2. Reglas lingüísticas y matrices asociadas

Finalmente, se definió una regla lingüística para cada categoría de operadores de la dimensión CUANTIFICACIÓN, a la que se le asignó su correspondiente matriz. A modo de ilustración, presentamos la regla y la matriz asociada (tabla 6) para la categoría ENG_CUANT_X4.

- (5) Regla lingüística de la categoría ENG_CUANT_X4: Si aparece un operador de esta categoría, antepuesto hasta en una posición a un adjetivo o pospuesto hasta en cinco posiciones a un verbo, con un foco cuya Polaridad de Entrada es positiva y su grado de intensidad es leve o fuerte, entonces el operador producirá un incremento de intensidad leve de la polaridad negativa en el caso de una Polaridad de Entrada positiva con un grado leve, y un incremento de intensidad fuerte de la polaridad negativa en el caso de una Polaridad de Entrada positiva con un grado fuerte.

TABLA 6

Matriz lingüística de la categoría ENG_CUANT_X4

| ETIQUETA | EXTENSIÓN ETIQUETA | | EXTENSIÓN ETIQUETA | | EXTENSIÓN ETIQUETA | | EXTENSIÓN ETIQUETA |
|-------------------------------------|--------------------|---|--------------------|---|--------------------|---|--------------------|
| <Cuant_Enunciado> | | | | | | | |
| ENG CUANT X4 | | | | | | | |
| ATRIBUTOS OPERADOR | | | | | | | |
| ALCANCE | 1 | | | | | | 5 |
| DIRECCIÓN | IZQ | | | | | | DCHA |
| SI | CASO 1 | | CASO 2 | | CASO 3 | | CASO 4 |
| ENG_CUANT_F4 | | | | | | | |
| ATRIBUTOS FOCO | | | | | | | |
| CATG | JJ | O | JJ | O | VB | O | VB |
| POLARIDAD_E | [POSITIVA, +1] | | [POSITIVA, +2] | | [POSITIVA, +1] | | [POSITIVA, +2] |
| ENTONCES | CASO 1 | | CASO 2 | | CASO 3 | | CASO 4 |
| POLARIDAD_MODIFICADOR | SÍ | | SÍ | | SÍ | | SÍ |
| T_POLARIDAD_MODIFICADOR | INTENSIFICADOR | | INTENSIFICADOR | | INTENSIFICADOR | | INTENSIFICADOR |
| CAMBIA_POLARIDAD | SÍ | | SÍ | | SÍ | | SÍ |
| INTENSIDAD_NIVEL | LEVE | | FUERTE | | LEVE | | FUERTE |
| FÓRMULA | MODIFICADA (9) | | MODIFICADA(10) | | MODIFICADA (9) | | MODIFICADA (10) |
| POLARIDAD_M | [NEGATIVA, -1] | | [NEGATIVA, -2] | | [NEGATIVA, -1] | | [NEGATIVA, -2] |
| <Cuant_Enunciado</> | | | | | | | |

5. Conclusiones y futuras líneas de investigación

En este artículo hemos propuesto un nuevo modelo computacional para tratar la dimensión de la cuantificación en inglés, dentro del reto que suponen los operadores contextuales de cambio de polaridad en la minería de opiniones. En concreto, hemos considerado que los operadores de cuantificación pueden modificar la polaridad inicial de una opinión incrementándola (Modificadores Intensificadores) o reduciéndola (Modificadores Atenuadores), en ambos casos con una intensidad leve o fuerte. Además, también hemos contemplado que no todos estos operadores contextuales pueden modificar la polaridad inicial de la misma manera y que, por ello, deben tratarse computacionalmente de forma específica.

En nuestro modelo, la detección automática y el tratamiento de estos operadores asociados a la cuantificación se realizan desde un enfoque simbólico, basado en el conocimiento, frente al enfoque estadístico, basado en el uso de técnicas de aprendizaje automático y corpus específicos. Al adoptar un enfoque simbólico, el modelo ofrece numerosas ventajas, ya que (a) no requiere de la creación de costosos corpus de entrenamiento, característicos de los sistemas estadísticos, (b) puede reutilizar distintos recursos ya existentes (e.g. lexicones de sentimiento y reglas lingüísticas) y (c) puede ser fácilmente replicado para el procesamiento de otras lenguas.

De este modo, hemos incorporado el conocimiento lingüístico teórico en una arquitectura modular, donde se agrupan los operadores en categorías según sus patrones de funciona-

miento, lo cual supone un notable avance con respecto a lo investigado hasta ahora, puesto que cada categoría recibe un tratamiento específico. Asimismo, hemos elaborado una serie de reglas lingüísticas asociadas a cada categoría que incluyen información sobre los operadores en cuanto, por un lado, al alcance y dirección con respecto a su foco y, por otro lado, a la intensidad del cambio que generan en la polaridad inicial a través del uso de fórmulas matemáticas. Finalmente, hemos diseñado unas matrices asociadas a estas reglas lingüísticas, las cuales representan instrucciones precisas para que este conocimiento lingüístico pueda ser fácilmente programable en un sistema computacional. Cabe destacar que este conocimiento lingüístico teórico sobre el que se basa la arquitectura propuesta ha sido contrastado con evidencias lingüísticas encontradas en los textos de los mensajes publicados en Twitter.

La arquitectura modular propuesta se incorporará a la base de conocimiento del proyecto ALLEGRO (Adaptive multi-domain social-media sEnsinG fRamework), un sistema para el desarrollo de aplicaciones multimodales que permitan reconstruir el estado de la sociedad desde el prisma de la inteligencia colectiva de los usuarios de redes sociales. En concreto, uno de sus módulos, DIAPASON (a unified hybrid Approach to microtext Analysis in Social-media cRowdseNsing), está diseñado para procesar contenido textual en inglés y español generado por usuarios con el fin de detectar automáticamente diversos tipos de problemas sociales integrando técnicas de procesamiento del lenguaje natural, aprendizaje automático e ingeniería del conocimiento. Por tanto, nuestro trabajo se centrará en la mejora y ampliación del modelo, extendiéndolo a otras lenguas e incorporando otras dimensiones de operadores contextuales.

6. Agradecimientos

Esta publicación es parte del proyecto de I+D+i PID2020-112827GB-I00, financiado por MICIU/AEI/10.13039/501100011033, y del proyecto SMARTLAGOON [101017861], financiado por Horizonte 2020 - Programa Marco de Investigación e Innovación de la Unión Europea.

7. Bibliografía citada

ÁLVAREZ, Isabel, 2011: "El ciberespañol: Características del español usado en internet", *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, 33-41.

ASGHAR, Muhammad Zubair, Aurangzeb KHAN, Syeda Rabail ZAHRA, Shakeel AHMAD y Fazal Masud KUNDI, 2017: "Aspect-based opinion mining framework using heuristic patterns", *Cluster Computing* 22, 7181-7199.

CARRILLO-DE-ALBORNOZ, Jorge, y Laura PLAZA, 2013: "An emotion-based model of negation, intensifiers, and modality for polarity and intensity classification", *Journal of the American Society for Information Science and Technology* 64 (8), 1618-1633.

CHEN, Xingming, Yanghui RAO, Haoran XIE, Fu Lee WANG, Yingchao ZHAO y Jian YIN, 2019: "Sentiment classification using negative and intensive sentiment supplement information", *Data Science and Engineering* 4 (2), 109-118.

CHEN, Boxing, y Xiaodan ZHU, 2014: "Bilingual sentiment consistency for statistical machine translation", *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 607-615.

CRUZ, Noa P., Maite TABOADA y Ruslan MITKOV, 2016: "A machine-learning approach to negation and speculation detection for sentiment analysis", *Journal of the Association for Information Science and Technology* 67 (9), 2118-2136.

ESCANDELL VIDAL, María Victoria, 2014: *La comunicación: Lengua, cognición y sociedad*, Madrid: Akal.

ESULI, Andrea, y Fabrizio SEBASTIANI, 2006: "SentiWordNet: A publicly available lexical resource for opinion mining", *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 417-422.

JIMÉNEZ-ZAFRA, Salud María, Roser MORANTE, María Teresa MARTÍN-VALDIVIA y L. Alfonso UREÑA-LÓPEZ, 2020: "Corpora annotated with negation: An overview", *Computational Linguistics* 46 (1), 1-52.

KENNEDY, Alistair, y Diana INKPEN, 2006: "Sentiment classification of movie reviews using contextual valence shifters", *Computational Intelligence* 22 (2), 110-125.

LI, Shoushan, Sophia Yat Mei LEE, Ying CHEN, Chu-Ren HUANG y Guodong ZHOU, 2010: "Sentiment classification and polarity shifting", *Proceedings of the 23rd International Conference on Computational Linguistics*, 635-643.

LI, Shoushan, Zhongqing WANG, Sophia Yat Mei LEE y Chu-Ren HUANG, 2013: "Sentiment classification with polarity shifting detection", *Proceedings of the 2013 International Conference on Asian Language Processing*, 129-132.

LIU, Bing, 2015: *Sentiment analysis: Mining opinions, sentiments, and emotions*, New York: Cambridge University Press.

LIU, Jingjing, y Stephanie SENEFF, 2009: "Review sentiment scoring via a parse-and-phrase paradigm", *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 161-169.

MACKENZIE, J. Lachlan, y Elena MARTÍNEZ-CARO, 2012: "Compare and contrast. An English grammar for speakers of Spanish" en Francisco SÁNCHEZ-BENEDITO (ed.): *Colección Estudios In-gleses*, vol. 20, Granada: Comares, 7-24.

MANCERA, Ana, y Ana PANO, 2013: *El español coloquial en las redes sociales*, Madrid: Arco Libros.

MARTIN, James R., y P. R. R. WHITE, 2005: *The language of evaluation: Appraisal in English*, London: Palgrave Macmillan.

MORENO-ORTIZ, Antonio, y Javier FERNÁNDEZ-CRUZ, 2015: "Identifying polarity in financial texts for sentiment analysis: A corpus-based approach", *Procedia - Social and Behavioral Sciences* 198, 330-338.

MORSY, Sara A., y Ahmed RAFAA, 2012: "Improving document-level sentiment classification using contextual valence shifters" en Gosse BOUMA, Ashwin ITTOO, Elisabeth MÉTAIS y Hans WORTMANN (eds.): *Natural language processing and information systems* (vol. 7337), Berlin-Heidelberg: Springer, 253-258.

MUDGAL, Pushpendra, y Ajay KHUNTETA, 2020: "Handling double intensifiers in feature-level sentiment analysis based on movie reviews" en Garima MATHUR, Harish SHARMA, Mahesh BUNDELE, Nilanjan DEY y Marcin PAPRZYCKI (eds.): *International Conference on Artificial Intelligence: Advances and Applications 2019. Algorithms for intelligent systems*, Singapore: Springer, 383-392.

MUSAT, Claudiu, y Stefan TRAUSSAN-MATU, 2010: "The impact of valence shifters on mining implicit economic opinions" en Darina DICHEVA y Danail DOCHEV (eds.): *Artificial intelligence: Methodology, systems, and applications. AIMSA 2010*, Berlin-Heidelberg: Springer, 131-140.

NEVIAROUSKAYA, Alena, Helmut PRENDINGER y Mitsuru ISHIZUKA, 2010: "@AM: Textual attitude analysis model", *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 80-88.

NEVIAROUSKAYA, Alena, Helmut PRENDINGER y Mitsuru ISHIZUKA, 2015: "Attitude Sensing in Text Based on A Compositional Linguistic Approach", *Computational Intelligence* 31 (2), 256-300.

NOVAK, Petra Kralj, Jasmina SMAILOVIĆ, Borut SLUBAN e Igor MOZETIČ, 2015: "Sentiment of emojis", *PloS One* 10 (12), 1-22.

PERIÑÁN-PASCUAL, Carlos, 2022: "Text mining" en Mark ARONOFF (ed.): *Oxford Bibliographies Online in Linguistics*, New York: Oxford University Press [<https://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0295.xml>, fecha de consulta: 25 de agosto de 2022].

PERIÑÁN-PASCUAL, Carlos, y Francisco ARCAS-TÚNEZ, 2019: "Detecting environmentally-related problems on Twitter", *Biosystems Engineering* 177, 31-48.

PHU, Vo Ngoc, y Phan Thi TUOI, 2014: "Sentiment classification using enhanced contextual valence shifters", *Proceedings of the 2014 International Conference on Asian Language Processing*, 224-229.

POLANYI, Livia, y Annie ZAENEN, 2004: "Contextual valence shifters", *Working Notes of the AAIL Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, 106-111.

QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH y Jan SVARTVIK, 1985: *A comprehensive English grammar*, London-New York: Longman.

RHEE, Seongha, 2016: "On the emergence of the stance-marking function of English adverbs: A case of intensifiers", *Linguistic Research* 33 (3), 395-436.

RUPPENHOFER, Josef, Jasper BRANDES, Petra STEINER y Michael WIEGAND, 2015: "Ordering adverbs by their scaling effect on adjective intensity", *Proceedings of the International Conference Recent Advances in Natural Language Processing*, 545-554.

SCHULDER, Marc, Michael WIEGAND y Josef RUPPENHOFER, 2020: "Enhancing a lexicon of polarity shifters through the supervised classification of shifting directions", *Proceedings of the 12th International Conference on Language Resources and Evaluation*, 5010-5016.

SINTSOVA, Valentina, Margarita BOLÍVAR JIMÉNEZ y Pearl PU, 2018: "Modeling the impact of modifiers on emotional statements" en Alexander GELBUKH (ed.): *Computational linguistics and intelligent text processing. CICLing 2017*, Cham: Springer, 71-89.

TABOADA, Maite, 2016: "Sentiment analysis: An overview from linguistics", *Annual Review of Linguistics* 2 (1), 325-347.

TABOADA, Maite, Julian BROOKE, Milan TOFILOSKI, Kimberly VOLL y Manfred STEDE, 2011: "Lexicon-based methods for sentiment analysis", *Computational Linguistics* 37 (2), 267-307.

VARATHAN, Kasturi Dewi, Anastasia GIACHANOU y Fabio CRESTANI, 2017: "Comparative opinion mining: A review", *Journal of the Association for Information Science & Technology* 68 (4), 811-830.

WHITELAW, Casey, Navendu GARG y Shlomo ARGAMON, 2005: "Using appraisal groups for sentiment analysis", *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, 625-631.

WILSON, Theresa, Janyce WIEBE y Paul HOFFMANN, 2009: "Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis", *Computational Linguistics* 35 (3), 399-433.

Yoo, Gwanghoon, y Jeesun NAM, 2018: "A hybrid approach to sentiment analysis enhanced by sentiment lexicons and polarity shifting devices" en *Proceedings of the 13th Workshop on Asian Language Resources*, 21-28.