

Minería de argumentación en el Referéndum del 1 de Octubre de 2017

Argument Mining in the October 1, 2017 Referendum

Marcos Esteve, Francisco Casacuberta, Paolo Rosso

Universitat Politècnica de València
maresca@inf.upv.es, {fcn, proso}@prhlt.upv.es

Resumen: La minería de argumentación permite, mediante herramientas software, obtener cuáles son los argumentos que expresan los autores en un determinado texto. En este artículo se pretende realizar un análisis de la argumentación expresada por los usuarios en Twitter en relación al referéndum del 1 de octubre de 2017. Se utilizará para ello, el dataset MultiStanceCat proporcionado en la tarea organizada en el IberEval 2018. Dado que las herramientas de minería de argumentación trabajan en su mayoría en inglés, será necesario construir un sistema de traducción neuronal con postedición que permita realizar una traducción de los tweets del español y el catalán al inglés. Los resultados al realizar la minería de argumentación sobre los *tweets* traducidos ha demostrado obtener un porcentaje muy reducido de argumentación en todas las comunidades.

Palabras clave: Minería de argumentación, traducción automática, 1Oct2017

Abstract: Argument mining allows, through software tools, to obtain which are the arguments expressed by the authors in a given text. This article aims to make an analysis of the arguments expressed by users on Twitter in relation to the referendum of October 1, 2017 using the MultiStanceCat dataset provided in the shared task organized at IberEval 2018. Since the tools of argumentation mining work mostly in English, it was necessary to build a neural translation system with post-editing that allows to translate of tweets from Spanish and Catalan to English. The results of argumentation mining on the translated tweets have shown to obtain a minimum percentage of argumentation in all the communities.

Keywords: Argument mining, machine translation, 1Oct2017

1 Introducción

Una de las principales problemáticas que existen en la actualidad en redes sociales, y que se acrecenta cuando los temas son controvertidos, como el referéndum sobre la independencia de Cataluña, el Brexit o las elecciones estadounidenses, es la alta polarización existente entre sus distintas comunidades (Lai, 2019; Lai et al., 2019). La polarización de las comunidades implica la existencia de una alta comunicación entre los usuarios de la misma comunidad y, además, los mensajes que intercambian con otras comunidades suelen ser mensajes tóxicos donde destaca el odio. En este tipo de comunidades existen algunos comportamientos propios de

la alta polarización existente. Por una parte en la comunicación dentro de la comunidad se puede observar el fenómeno de aislamiento, donde el usuario solo percibe información perteneciente a su comunidad viendo de esta forma una sola realidad. Relacionado con el anterior, también se observa el fenómeno de cámara de eco, donde, debido al aislamiento del usuario se produce un incremento de sus creencias.

El objetivo principal de este artículo consiste en determinar si existe argumentación en los *tweets* de la tarea MultiStanceCat (Taulé et al., 2018) presentada en el IberEval del 2018. Este dataset consta de un conjunto de *tweets* agrupados según su posicionamiento respecto al referéndum del 1 de octubre de

2018, a favor, en contra, o neutral. Para realizar el análisis, será necesario hacer uso de herramientas de minería de argumentación, que permitan analizar dichos *tweets* extrayendo los argumentos expresados por los usuarios de las distintas comunidades. Dado que no existen tecnologías desarrolladas de minería de argumentación para *tweets* en español o catalán, se ha optado por utilizar la herramienta desarrollada por Iryna Gurevych en la Universidad Técnica de Darmstadt¹. Esta herramienta únicamente trabaja en inglés o alemán, por lo que será necesario traducir los *tweets* a dicha lengua. Para ello se ha propuesto un sistema de traducción con postedición basado en redes neuronales, con el que facilitar al operario humano la tarea de traducción de dichos *tweets*.

En los siguientes apartados se detallará el estado del arte para las distintas tecnologías y una introducción a los datasets utilizados. Se comentarán los sistemas desarrollados, así como la experimentación realizada para la traducción neuronal. Por último, se detallarán los experimentos desarrollados para la minería de argumentos sobre los *tweets* traducidos.

2 Estado del arte

2.1 Minería de argumentación

La argumentación permite a un escritor convencer a una audiencia sobre unas ideas. En esta área, un argumento es una estructura que consta de varios componentes donde, tal y como comentan los autores en (Stab y Gurevych, 2014), se fundamenta por una afirmación que está apoyada o atacada por, al menos, una premisa. La afirmación tomará por tanto, la parte principal del argumento, y le seguirán las premisas que trataran de sustentar la afirmación, persuadiendo de esta forma al lector. Por ejemplo, tal como detallan los autores y se puede observar en la Figura 1, la afirmación **1** tomará la parte principal del argumento y le seguirán las premisas **2** y **3**, que tratarán de sustentar la afirmación. Por último, la premisa **4** sustentará a su vez a la premisa **3**.

“(1) Museums and art galleries provide a better understanding about arts than Internet. (2) In most museums and art galleries, detailed descriptions in terms of the background, history and author are provided. (3) Seeing an artwork online is not the same as watching it with our own eyes, as (4) the picture online does not show the texture or three-dimensional structure of the art, which is important to study.”

Figura 1: Ejemplo de estructura argumentativa

La minería de argumentación permite, de manera automática, extraer estructuras argumentativas de grandes volúmenes de textos. Para ello, tal como indican los autores en (Eger, Daxenberger, y Gurevych, 2017), las investigaciones se centran en varias sub-tareas:

- La separación de unidades argumentativas frente a no argumentativas (segmentación en componentes)
- La clasificación de componentes argumentativas en premisa o afirmación
- Búsqueda de relaciones entre las estructuras argumentativas
- Clasificación de las relaciones según apoyen o ataquen a la afirmación

Un aspecto interesante a destacar es la transición existente en el desarrollo de dichas tecnologías. Mientras que, históricamente, la búsqueda de estructuras argumentativas se basaba en el uso de gramáticas construidas manualmente para un dominio específico (Palau y Moens, 2009), en la actualidad las tareas de minería de argumentación se están centrando en sistemas *end-to-end* basados en redes neuronales recurrentes (Eger, Daxenberger, y Gurevych, 2017).

2.2 Traducción automática

La traducción automática trata de solventar el problema de la traducción de una lengua origen a otra destino, utilizando para ello herramientas automáticas. Esta área ha sufrido un alto auge gracias al incremento de textos bilingües paralelos donde, para cada oración en un idioma origen, se dispone de su correspondiente traducción en una lengua destino.

¹<http://www.argumentsearch.com/>

Históricamente, la traducción automática estaba basada en el conocimiento lingüístico, hasta la irrupción con gran éxito, de la llamada traducción estadística (Brown et al., 1990; Koehn y Knight, 2000), donde cada oración origen podía ser traducción de una oración destino y, por tanto, el objetivo consistía en asignar una alta probabilidad en caso de que, efectivamente, fueran traducción una de otra y, en caso contrario, asignarle una baja probabilidad. En la ecuación 1 se detalla la fórmula básica de la traducción estadística, donde x es la oración a traducir e y es una posible traducción.

$$\operatorname{argmax}_{I,s} \prod_{i=1}^I p(y_i | y_1^{i-1}, x) \quad (1)$$

Actualmente, los esfuerzos en traducción se están centrando en la traducción neuronal, ya que ha demostrado obtener mejores resultados frente a la traducción estadística (Koehn, 2009). Esto ha sido gracias, en parte, a la inclusión de las representaciones distribucionales conocidas como *word embeddings* (Bengio et al., 2003; Mikolov et al., 2013). Este tipo de modelado permitió representar las palabras como vectores densos que capturaban las relaciones semánticas y sintácticas con otros términos. Gracias a esto último ha sido posible aplicar potentes técnicas de redes neuronales a una gran cantidad de tareas textuales, obteniendo resultados del actual estado del arte. En el caso de traducción automática, las dos arquitecturas que fijaron el actual estado del arte, son por una parte, la arquitectura encoder-decoder con modelo de atención (Luong, Pham, y Manning, 2015) basada en redes neuronales recurrentes y, por otra parte, y más recientemente el uso de modelos basados en Transformers (Vaswani et al., 2017); basados en redes *feed-forward* y modelos de auto-atención entre las palabras de entrada, las palabras de salida y las palabras de entrada y salida.

Por otra parte, dado que los resultados en traducción siguen sin permitir una traducción totalmente automática existen numerosos esfuerzos por conseguir sistemas que permitan a un operador experto traducir texto con el mínimo esfuerzo posible (Peris, Domingo, y Casacuberta, 2017). En esta área se han desarrollado numerosas técnicas de postedición con las que se pueden obtener mejores resultados con un esfuerzo humano inferior a

si se tuviera que traducir manualmente.

3 Datasets

Dado que no existen herramientas para la extracción de argumentos que trabajen en español o catalán y, que la herramienta desarrollada por la Universidad Técnica de Darmstadt solo trabaja en inglés o alemán; se ha decidido integrar un sistema de traducción que permita traducir los *tweets* proporcionados en el dataset MultiStanceCat del español y el catalán al inglés. Para ello se requiere de corpus paralelos que permitan la obtención de traductores tanto para el catalán como para el español.

3.1 Dataset MultiStanceCat

El corpus MultiStanceCat (Taulé et al., 2018) consiste en una recopilación de *tweets* multimodal (texto e imagen), utilizando para ello los *hashtags* #1oct, #1O, #1oct2017 y #1oct16. Tras la recopilación de 220.148 *tweets* entre el español y el catalán, los autores anotaron manualmente un subconjunto de estos, dependiendo de su posicionamiento respecto al referéndum (a favor, en contra o neutral) obteniendo de esta forma un total de 5853 *tweets* en catalán y 5545 en español. Además, con el fin de garantizar la calidad de las anotaciones, los autores realizaron dos procesos de acuerdo inter-anotador.

Posicionamiento	Catalán	Español
A favor	5106	2099
En contra	149	2231
Neutral	598	1215
Total	5853	5545

Tabla 1: Distribución de los *tweets* según idioma y según se posicionan a favor, en contra o neutral

El número de datos recopilados está balanceado en cuanto a número de *tweets* entre ambas lenguas. Además, tal y como se observa en la tabla, mientras para el español la distribución de *tweets* en las distintas clases está balanceada, para el catalán se observa un sesgo donde la mayoría de los *tweets* recopilados se han etiquetado con un posicionamiento positivo.

Por otra parte, los autores realizaron un estudio con el fin de determinar cuanto polarizadas estaban las distintas comunidades respecto al referéndum. El estudio determi-

naba que las comunidades estaban desconectadas entre si, ya que únicamente el 13.64% de los usuarios seguían a usuarios de otras comunidades y, el número de usuarios que solo seguían a usuarios con una polarización, positiva o negativa, era del 51.44% y 25.03% respectivamente. Además, únicamente el 9.89% de los usuarios manifestaban una postura neutral, por lo que existía una alta polarización.

3.2 Datasets para la traducción automática

Para la construcción de los traductores ha sido necesario el uso de corpus paralelos de oraciones español-inglés y catalán-inglés. Se han utilizado el corpus Europarl para traducir del español al inglés y el Opensubtitles para traducir del catalán al inglés.

3.2.1 Europarl

El Europarl (Tiedemann, 2012) se trata de un corpus paralelo extraído de la web del Parlamento Europeo por Philipp Koehn. Este corpus incluye versiones en 21 lenguas europeas. Destacar que el corpus que nos interesa consta de un total de 1.9 millones de pares de frases español-inglés y alrededor de 50 millones de palabras, tanto para el español como para el inglés. A estos pares será necesario aplicarles un preproceso, así como una tokenización.

3.2.2 OpenSubtitles

El corpus OpenSubtitles (Lison y Tiedemann, 2016) se trata de un corpus paralelo extraído de subtítulos de películas. Este corpus ha sido extraído de la pagina web <http://www.opensubtitles.org/>. El corpus incluye versiones para 62 lenguas distintas. La versión catalán-inglés consta de un total de 428.000 pares de oraciones y alrededor de 3 millones de palabras. Estos pares catalán-inglés deberán ser preprocesados y tokenizados para poder ser utilizados por el traductor.

4 Traducción neuronal

4.1 Propuesta de solución

A la hora de construir los traductores se ha propuesto un sistema basado en redes neuronales, utilizando para ello el *toolkit* NMT-Keras (Peris y Casacuberta, 2018). Dado que los traductores que se van a entrenar están basados en corpus con un lenguaje más formal y el objetivo final es obtener traducciones de un corpus de *tweets*, los cuales poseen un

lenguaje coloquial lleno de errores, abreviaturas etc. Se ha propuesto un sistema basado en postedición que permita a un humano experto traducir de una manera eficiente e ir mejorando las prestaciones de los traductores de forma iterativa (Domingo et al., 2019; Peris y Casacuberta, 2019). En la figura 2 se detalla la implementación seguida para dicho sistema.

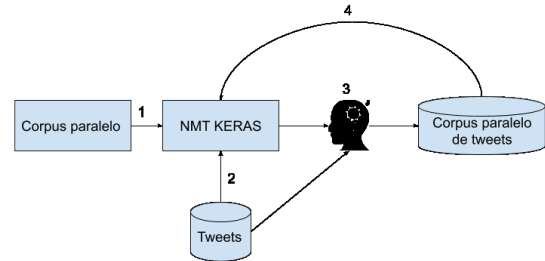


Figura 2: Descripción del sistema utilizado

Tal y como se detalla en la figura 2, el sistema consta de distintas fases:

1. Entrenamiento de los traductores neuronales con el *toolkit* NMT-Keras y utilizando los corpus paralelos expuestos en 3.2.1 y 3.2.2.
2. Extracción de un subconjunto de *tweets* del dataset expuesto en 3.1 y traducción por el sistema entrenado.
3. Revisión por parte del operario experto de las traducciones generadas por el sistema, rectificándolas y generando de esta forma un corpus paralelo de *tweets*.
4. Ajuste de los parámetros del traductor neuronal, utilizando el corpus paralelo de *tweets* generado por el operario.

4.2 Experimentación

Para realizar la experimentación se ha cogido un subconjunto de los datasets paralelos y se ha explorado la variación del BLEU al cambiar los parámetros entre los modelos de **encoder-decoder con atención y transformer**. Concretamente, para las pruebas se ha utilizado 180.000 muestras para entrenamiento, 1.000 muestras para desarrollo y 2.500 muestras para test.

4.2.1 Resultados Europarl

En primer lugar, se ha explorado cual es la contribución del tamaño del *embedding* en el modelo de *encoder-decoder*; para ello se ha fijado el tamaño de las redes LSTM a 64 unidades y se han variado los tamaños de los

embeddings, tanto para el *encoder* como para el *decoder*.

Input	Target	BLEU
64	64	22.81
128	128	30.99
256	256	30.01

Tabla 2: Evolución del BLEU al modificar el tamaño de los embeddings en el corpus Europarl (3 epochs)

Como se puede observar en la tabla 2 los mejores resultados se obtienen cuando el tamaño de *embedding* es 128; por esta razón se fijará el tamaño del *embedding* a este tamaño a la hora de realizar la experimentación, y así observar la contribución del tamaño de la red LSTM a la variación en el BLEU.

Input LSTM	Target LSTM	BLEU
64	64	30.99
128	128	32.12
256	256	31.43

Tabla 3: Evolución del BLEU al modificar el tamaño de las capas LSTM en el corpus Europarl (3 epochs)

Tal y como se puede observar en la tabla 3, para nuestro caso el mejor modelo basado en *encoder-decoder* con modelo de atención se obtiene cuando el tamaño de los *embeddings* y de la red coinciden con 128 unidades.

Por otra parte, se ha explorado la arquitectura de traducción basada en Transformers, aunque las experimentaciones han revelado que los resultados obtenidos son significativamente inferiores a los obtenidos con la arquitectura *encoder-decoder*.

Por último, una vez determinada la mejor arquitectura (*encoder-decoder*) y parametrización (tamaño de *embeddings* 128 y unidades LSTM 128) se ha entrenado un traductor con un mayor volumen de datos, obteniendo un BLEU de 33.91. Este será, por tanto, el traductor que traducirá los *tweets* y se irá mejorando de forma iterativa utilizando el sistema de postedición propuesto.

4.2.2 Resultados OpenSubtitles

Se ha seguido un enfoque similar al expuesto anteriormente, fijando primeramente para la arquitectura *encoder-decoder* el tamaño de la red LSTM a 64 unidades y explorando la variación del BLEU al variar el tamaño del *embedding*.

Input	Target	BLEU
64	64	22.28
128	128	22.59
256	256	21.61

Tabla 4: Evolución del BLEU al modificar el tamaño de los embeddings en el corpus OpenSubtitles (3 epochs)

A la vista de los resultados expuestos en la tabla 4, el mejor BLEU se obtiene cuando el tamaño de *embedding* de entrada y salida coincide con 128 unidades. Se fijará, por tanto, este tamaño a los *embedding* a la hora de explorar la evolución del BLEU al aumentar el número de unidades en las redes LSTM.

Input LSTM	Target LSTM	BLEU
64	64	22.59
128	128	23.73
256	256	23.91
512	512	22.71

Tabla 5: Evolución del BLEU al modificar el tamaño de las redes LSTM en el corpus OpenSubtitles (3 epochs)

Como se puede observar en la tabla 5, los mejores resultados se obtienen cuando el tamaño de *embedding* es de 128 y el tamaño de las redes LSTM coincide con 256 unidades.

Además de la arquitectura *encoder-decoder*, se ha explorado la variación del BLEU al modificar el tamaño del modelo basado en la arquitectura Transformer, aunque los resultados no han conseguido mejorar al modelo basado en la arquitectura *encoder-decoder*.

De nuevo, una vez determinada la mejor arquitectura (*encoder-decoder*) y parametrización (tamaño de *embeddings* 128 y LSTM 256), se ha entrenado un traductor con un mayor volumen de datos obteniendo un BLEU de 27.02. Este será el nuevo traductor neuronal (catalán-inglés), el cual se irá mejorando de forma progresiva en la traducción de *tweets* empleando el sistema de postedición propuesto.

4.2.3 Traducción de los tweets

Una vez entrenados los traductores neuronales para la traducción del español y el catalán al inglés, se ha seguido la propuesta comentada en el apartado 4.1 para traducir los *tweets*. Se observó un incremento de las prestaciones de los traductores neuronales al incrementar

el número de muestras posteditadas en el proceso de adaptación; obteniendo traducciones más eficientes a si los tweets se tuvieran que traducir de forma manual.

5 Argument mining

Una vez traducido el corpus de *tweets* del dataset comentado en el apartado 3.1, se ha hecho uso de la herramienta, desarrollada por la Universidad Técnica de Darmstadt, ArgumentText (Stab et al., 2018) y disponible en el enlace <http://www.argumentsearch.com/>. Se trata de una herramienta que permite buscar argumentos escritos en inglés o alemán en grandes colecciones de documentos, haciendo uso de aprendizaje automático. Más concretamente hace uso de redes neuronales con modelos de atención para la clasificación de las oraciones según si poseen argumentos o no y, posteriormente redes recurrentes BILSTM, para clasificar el posicionamiento a favor o en contra del tema que está siendo analizado. Dado que la herramienta utilizada necesita un tema sobre el que realizar el análisis, se ha utilizado la palabra clave *catalonia* y se han agrupado los *tweets*, dependiendo de su lengua origen y del posicionamiento que tuvieran en el corpus MultiStanceCat.

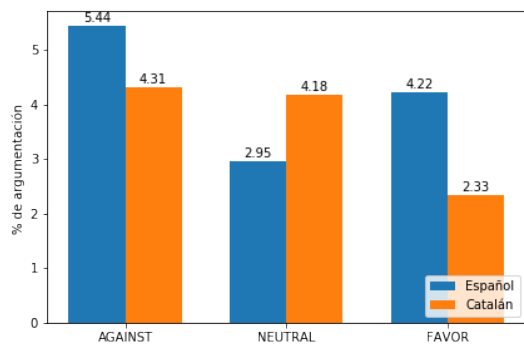


Figura 3: Porcentaje de argumentación según el idioma y el posicionamiento

En la figura 3 se puede observar como el porcentaje de argumentación es muy bajo, máximo de 5.44% en el caso de la comunidad que está en contra y habla en español. Además, se puede observar que los hablantes de español en el caso de las comunidades a favor y en contra del referéndum, tratan de argumentar en mayor medida su postura, en comparación con los hablantes de catalán. En cambio, en el caso de la comunidad neutral se observa como los hablantes de catalán son los que tratan de argumentar en mayor medida

sus ideas.

Algunos de los argumentos encontrados por la herramienta son:

- El de @cupnacional diu que tindrem més drets indepes. Potser si,BarcelonaWorld i escoles OPUS De moment recolzament partit corrupte. #1octL6² (**En contra**)
- #1O Catalá diu que especular amb la detenció de Puigdemont té poc sentit i insisteix que és decisió del jutge³ (**Neutral**)
- #1octL6 que nivel de irresponsabilidad solo por interes electoral, de unos y otros. Politicos rompiendo la sociedad y un pais. Que pena...(**En contra**)
- Dice el Gobierno español q si renunciamos al referéndum #1O nos darán más dinero, autonomía y reformarán la Constit... (**A favor**)

6 Conclusiones y trabajo futuro

El objetivo principal de este artículo consistía en determinar si existía argumentación en los *tweets* del dataset MultiStanceCat acerca del Referéndum del 1 de Octubre de 2017 sobre la independencia de Cataluña. Para ello, puesto que no existen herramientas que realicen minería de argumentos sobre el catalán o el español, en este trabajo se ha propuesto un sistema de traducción neuronal construido con en el *toolkit* Nmt-Keras y basado en postedición para traducir el corpus de *tweets* de la tarea MultiStanceCat del catalán y el español al inglés. Esto ha permitido adaptar un dominio formal, como pueda ser un traductor entrenado sobre el corpus Europarl a la traducción de un dominio informal lleno de errores, como pueda ser los *tweets*. Esta fase ha sido especialmente ardua debido a la propia jerga informal utilizada en Twitter, donde por lo general existen una gran cantidad de errores en la escritura, abreviaciones, coetillas y características propias de la comunicación espontánea en medios sociales. Para solucionar este problema, ha sido necesario realizar un gran número de post-ediciones con el fin de

²El de @cupnacional dice que tendremos más derechos Indep. Quizás si, BarcelonaWorld y escuelas OPUS de momento apoyemos al partido corrupto. #1octL6

³#1O Catalá dice que especular con la detención de Puigdemont tiene poco sentido e insiste en que es decisión del juez

adaptar los traductores iniciales a la traducción de los *tweets*.

Por otra parte, una vez traducidos los *tweets* y, a la hora de analizar la argumentación existente en el corpus MultiStanceCat, se ha observado que los usuarios, en su mayoría no tratan de argumentar su postura acerca del referéndum celebrado el 1 de Octubre de 2018 y, además, existe un pequeño incremento en la argumentación de los usuarios castellanoparlantes.

Como trabajo futuro, puede ser interesante realizar un estudio del origen de los errores, ya que pueden existir errores en la traducción de los tuits y en la extracción de los argumentos. Por último, también puede ser interesante estudiar cuál es el porcentaje de argumentación en un corpus de tweets genérico. Esto permitiría realizar mejores comparaciones, permitiendo determinar si es habitual el uso de argumentación en Twitter. Además, se puede destacar la necesidad que existe de desarrollar herramientas que permitan realizar minería de argumentación para el español y las demás lenguas del estado español. Esto permitiría obtener mejores resultados al no tener que aplicar técnicas de traducción automática.

Agradecimientos

Dicho trabajo ha sido desarrollado en el marco de la asignatura Traducción Automática del Máster Universitario en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital (MIARFID) de la Universitat Politècnica de València. El trabajo de los últimos dos autores se ha desarrollado en el marco del proyecto *Misinformation and Miscommunication in social media: FAKE news and HATE speech* (MISMIS-FAKEHATE) del Ministerio de Ciencia, Innovación y Universidades (PGC2018-096212-B-C31) y del proyecto PROMETEO/2019/121 (DeepPattern) de la Generalitat Valenciana.

Bibliografía

- Bengio, Y., R. Ducharme, P. Vincent, y C. Jauvin. 2003. A neural probabilistic language model. *Journal of machine learning research*, 3:1137–1155.
- Brown, P. F., J. Cocke, S. A. Della Pietra, V. J. Della Pietra, F. Jelinek, J. Lafferty, R. L. Mercer, y P. S. Roossin. 1990. A statistical approach to machine translation. *Computational linguistics*, 16(2):79–85.
- Domingo, M., M. García-Martínez, Á. Peris, A. Helle, A. Estela, L. Bié, F. Casacuberta, y M. Herranz. 2019. Incremental adaptation of nmt for professional post-editors: A user study. *arXiv preprint arXiv:1906.08996*.
- Eger, S., J. Daxenberger, y I. Gurevych. 2017. Neural end-to-end learning for computational argumentation mining. En *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*.
- Koehn, P. 2009. *Statistical machine translation*. Cambridge University Press.
- Koehn, P. y K. Knight. 2000. Estimating word translation probabilities from unrelated monolingual corpora using the em algorithm. En *AAAI/IAAI*, páginas 711–715.
- Lai, M. 2019. *On language and structure in polarized communities*. Ph.D. tesis, Universitat Politècnica de València.
- Lai, M., M. Tambuscio, V. Patti, G. Ruffo, y P. Rosso. 2019. Stance polarity in political debates: A diachronic perspective of network homophily and conversations on twitter. *Data & Knowledge Engineering*, 124 (online).
- Lison, P. y J. Tiedemann. 2016. Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles.
- Luong, M.-T., H. Pham, y C. D. Manning. 2015. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*.
- Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, y J. Dean. 2013. Distributed representations of words and phrases and their compositionality. En *Advances in neural information processing systems*, páginas 3111–3119.
- Palau, R. M. y M.-F. Moens. 2009. Argumentation mining: the detection, classification and structure of arguments in text. En *Proceedings of the 12th international conference on artificial intelligence and law*, páginas 98–107.

- Peris, Á. y F. Casacuberta. 2018. NMT-Keras: a very flexible toolkit with a focus on interactive NMT and online learning. *The Prague Bulletin of Mathematical Linguistics*, 111:113–124.
- Peris, Á. y F. Casacuberta. 2019. Online learning for effort reduction in interactive neural machine translation. *Computer Speech & Language*, 58:98–126.
- Peris, Á., M. Domingo, y F. Casacuberta. 2017. Interactive neural machine translation. *Computer Speech & Language*, 45:201–220.
- Stab, C., J. Daxenberger, C. Stahlhut, T. Miller, B. Schiller, C. Tauchmann, S. Eger, y I. Gurevych. 2018. ArgumenText: Searching for arguments in heterogeneous sources. En *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*.
- Stab, C. y I. Gurevych. 2014. Identifying argumentative discourse structures in persuasive essays. En *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, páginas 46–56.
- Taulé, M., F. M. R. Pardo, M. A. Martí, y P. Rosso. 2018. Overview of the task on multimodal stance detection in tweets on catalan# 1oct referendum. En *IberEval@SEPLN*, páginas 149–166.
- Tiedemann, J. 2012. Parallel data, tools and interfaces in opus. En *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*.
- Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, y I. Polosukhin. 2017. Attention is all you need. En *Advances in neural information processing systems*, páginas 5998–6008.