

Emotion Mining Through The Traductological Study Of Documents Most Downloaded And Read News

Sir-Alexci Suárez Castrillón¹, July A. Gómez-Camperos², Haidee Y. Jaramillo³

¹Department of Systems Engineering, Faculty of Engineering, GRUCITE Research Group, University Francisco de Paula Santander Ocaña, Colombia

²Mechanical Engineering Department, Faculty of Engineering, GINSTI Research Group, University Francisco de Paula Santander Ocaña, Colombia

³Department of Civil Engineering, Faculty of Engineering, GINSTI Research Group, University Francisco de Paula Santander Ocaña, Colombia

ABSTRACT

The research determines the change of emotions that may occur when translating a Spanish document into English, which may diminish or have repercussions and not be favorable to the official language. The analyzed documents are obtained from three databases: the first one of engineering documents, the second one of finance and the third one is obtained from the news archive with the current information topics. A sample of three documents is obtained for each database, from the most downloaded files and most read news. The difference between the values of the change of emotions is obtained, determining that if the difference is less or greater than 5%, it does not affect the analysis. The results show that when translating them the emotions are more neutral and decrease with respect to the official language, therefore the appropriate vocabulary should be analyzed to maintain the value or magnitude of the original feelings.

Keywords: emotion mining; document translation; artificial intelligence; news history.

1. INTRODUCTION

Emotions move markets, companies and decisions in many aspects of people's daily lives. The analysis can be done by reviewing the emotions found in the texts or opinions of users, as well as the textual messages of many companies, even scientific papers can be read in their entirety according to the perception that the reader captures from the text, where sometimes the analysis that is performed is known as emotion mining.

Emotions however do not have the same meaning in all languages or at least may differ from the official language in which it is written (Criado, 2019), even the change of words in a translation can generate better emotions or weaken the message (nuteco, 2019). The study by Saralegui Mavrou (Mavrou & Saralegui, 2018), conduct a study in three different languages with 40 participants, where the complexity of emotions in each language is discussed, including being the participants bilingual and multilingual speakers, where the lack of emotional

expression in some cases is due to lack of a broader vocabulary. A translational study analyzes emotions in Spanish/English financial reports (Orts, 2020), where it was found that the Spanish text was generally longer than the English text, but no significant change was observed, where intensity and polarity are preserved in the original and translated text.

The objective of this research is to try to determine if the translation of scientific documents and news of the Universidad Francisco de Paula Santander Ocaña (UFPSO), could differ from the original when translated into English, due to the importance it may have in processes such as the office of international relations (ORI), to support internationality, and that more foreign students enter the university or in scientific documents to be read and cited by researchers in English. For this purpose, two scientific databases are analyzed, the first one on engineering and the second one on finance, and a third one with the history of news published by UFPSO on a daily basis. The most downloaded scientific documents and most read news are analyzed, within the most downloaded documents in June 2022, is the research of Carrascal (2011), where he shows the use of polycubes as didactic tools that help the learning of mathematics, to understand the basic concepts of area, volume and perimeter, the research is supported with the implementation in the Educational Institution Artistic School Rafael Contreras Navarro of Ocaña. The work presenting the pedagogical models used at UFPSO (Durán-Chinchilla et al., 2016), aims to determine whether the practices carried out by teachers at the university are in line with the pedagogical model of Social Constructivism, so that they can be aligned according to institutional guidelines. Also within the most downloaded documents is the design to build a mill (Florez-Solano et al., 2014), to produce glue for ceramic products, where the importance is highlighted because it is one of the star products of manufacture and export of the Norte de Santander region.

Within the area of finance, the most downloaded documents are organizational theory (Gutiérrez et al., 2021), as an initiative to improve production in companies in order to achieve collective welfare rather than individual welfare. The postulates by Taylor and Fayol are analyzed (Agámez & Gúzman, 2015), as rules that are still being applied in modern business development, after almost 6 decades of being made known. Also among the most downloaded is the study on the use and coverage of ICT in rural areas of Colombia (Carbonell et al., 2020), its importance lies in knowing the coverage and the possibility that these tools are available in areas that in addition to being rural present an inconvenience by territory of difficult access, and the study is delimited on the educational institutions located in the areas of analysis.

Among the most read news on the UFPSO web page and the news archive for the month of June 2022, is the participation of the UFPSO in the meeting of postgraduate coordinators (UFPSOa, 2022), as well as the teaching of courses such as ArcGis, which allows the introduction to geographic information systems (UFPSOb, 2022). recognition of teaching work (UFPSOc, 2022) also becomes one of the most read news items for that period, based on 4 indicators: participation in research events, higher scores in student evaluation, participation in events organized by the department and participation in academic teaching mobility.

2. METHODOLOGY

In order to analyze whether translating scientific documents or UFPSO news into another language can generate a change of emotions, which can affect the context of the text and its possible complete reading of the information, the original documents in Spanish and their translation into English are used. The totality of scientific documents are from 2 databases and from the web page of the historical news registry for the month of June 2022. From each of the databases, the 3 most downloaded documents and also 3 news items within the historical record are selected, making a total of 9 documents to be analyzed. The software used is Wolfram, which allows us to compare each of the documents and their corresponding translation. Figure 1 shows the process used to determine the analysis of emotions.

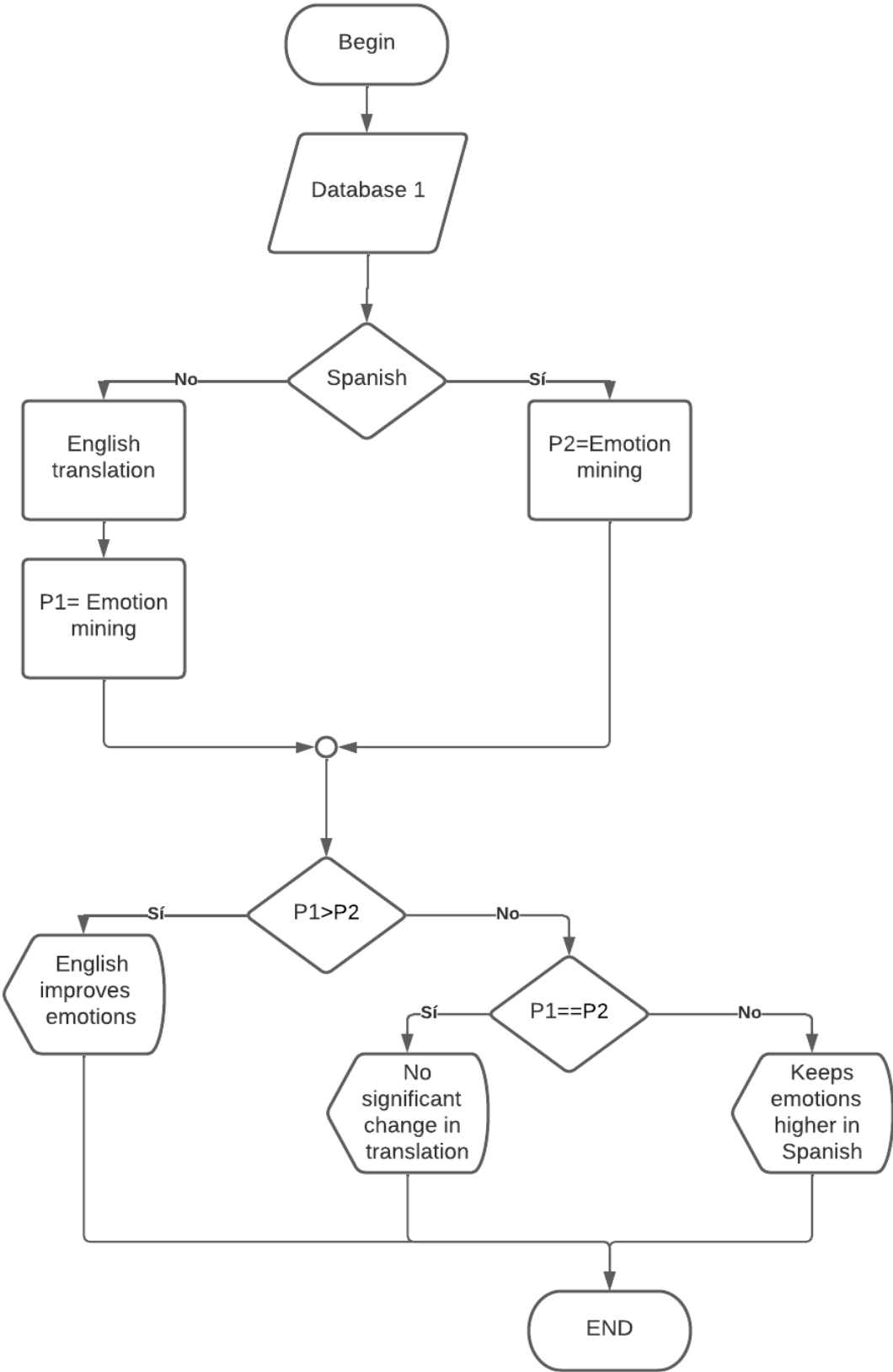


Figure 1. Emotion evaluation process.

Two values p_1 for English and p_2 for Spanish are obtained from each document to determine if the change is significant; the value does not necessarily have to be the same for there to be no change, a 5% error is taken as a reference to maintain equality, based on the sensitivity equations.

$$P1 = \sum_{DE=1}^3 DE * \left(\frac{TP}{TP + FN} \right) \quad (1)$$

$$P2 = \sum_{DS=1}^3 DS * \left(\frac{TP}{TP + FN} \right) \quad (2)$$

We show 3 examples from each of the databases, where it can be seen that the English texts will always become shorter than the English language, starting with the work of polycubes (Carrascal-Carrascal, 2011) in Spanish:

“El uso de policubos como recurso didáctico en el aula de matemáticas es ampliamente documentado. Un policubo es un agregado de un número determinado de cubos de igual tamaño, unidos por una de sus caras. Estos permiten la construcción de conceptos relacionados con perímetros, áreas y volúmenes, el desarrollo de la capacidad para resolver problemas, la habilidad para una búsqueda constante de propuestas novedosas relacionadas con un área específica del conocimiento, además de potenciar habilidades de investigación enmarcadas en un ambiente colaborativo. Un caso particular de los policubos es el conocido Cubo de Soma (Serrentino, 2011), ideado por el matemático Piet Hein, el cual consta de siete policubos irregulares formados por cuatro o menos cubos de igual tamaño (para un total de 27 cubos). El presente artículo detalla la experiencia desarrollada con los estudiantes de la básica secundaria y media técnica de la Institución Educativa Colegio Artístico Rafael Contreras Navarro de Ocaña, Norte de Santander, desde el año 2007 hasta la fecha; la cual busca el desarrollo y fortalecimiento del pensamiento lógico y la intuición espacial a través de la construcción y manipulación del cubo en mención.”

English translation:

“The use of polycubes as a didactic resource in the mathematics classroom is widely documented. A polycube is an aggregate of a determined number of cubes of equal size, joined by one of their faces. These allow the construction of concepts related to perimeters, areas and volumes, the development of problem solving skills, the ability to constantly search for novel proposals related to a specific area of knowledge, as well as to enhance research skills in a collaborative environment. A particular case of polycubes is the well-known Soma Cube (Serrentino, 2011), devised by the mathematician Piet Hein, which consists of seven irregular polycubes formed by four or fewer cubes of

equal size (for a total of 27 cubes). This article details the experience developed with high school and middle school students of the Institución Educativa Colegio Artístico Rafael Contreras Navarro de Ocaña, Norte de Santander, from 2007 to date, which seeks the development and strengthening of logical thinking and spatial intuition through the construction and manipulation of the cube in question.”

From the Spanish organization development (Gutiérrez et al., 2021):

“El desarrollo organizacional está relacionado con la calidad de vida de los trabajadores y cómo ello puede contribuir a mejoras en una unidad empresarial, ese es el punto al que se ha dado importancia desde hace tanto tiempo: el clima laboral ya que es un factor de productividad que puede funcionar como indicador de rendimiento en cualquier empresa. El presente artículo se enfoca en la investigación de estrategias de organización que contribuyan al desarrollo laboral y administrativo con un fin de bienestar colectivo en la empresa. Se realizan búsquedas documentales y observaciones sobre los métodos organizativos y el desarrollo humano en Bavaria S.A como caso de estudio, para conocer el rendimiento interno y externo en relación a la teoría del desarrollo organizacional (DO), con base a esta sus ventajas y estrategias de implementación, vigencia, y cómo la teoría misma permite alcanzar altos niveles de competitividad y posicionamiento en cualquier tipo de mercado dando como resultado la ejemplificación de una empresa que en la actualidad se destaca por su rendimiento en relación al ámbito laboral, productivo y administrativo con directrices organizacionales.”

Its English translation:

“Organizational development is related to the quality of life of workers and how this can contribute to improvements in a business unit, that is the point to which importance has been given for so long: the work environment, since it is a productivity factor that can function as a performance indicator in any company. This article focuses on the investigation of organizational strategies that contribute to labor and administrative development with the aim of collective welfare in the company. Documentary searches and observations on organizational methods and human development in Bavaria S.A. as a case study are made, to know the internal and external performance in relation to the theory of organizational development (OD), based on this its advantages and implementation strategies, validity, and how the theory itself allows reaching high levels of competitiveness and positioning in any type of market resulting in the exemplification of a company that currently stands out for its performance in relation to the labor, productive and administrative field with organizational guidelines.”

And the first within the most named in Spanish taken from the official website for the month of June (UFPSOa, 2022), in Spanish:

“La Red Colombiana de Posgrados programó el Primer Encuentro de Coordinadores de Posgrado por área del Conocimiento, desarrollado el 15 de junio de 2022 en modalidad híbrida (Presencial y Virtual) y siendo dirigido a coordinadores de posgrado de las

instituciones que hacen parte de la Red Colombiana de Posgrados. El objetivo del evento fue dar a conocer experiencias y buenas prácticas en la gestión académica y administrativa de los posgrados ofrecidos por las IES que integran la Red, en distintas áreas del conocimiento. Durante el evento se llevaron a cabo dos conferencias El mundo “analógico” del pasado, frente al mundo “inteligente” del futuro, a cargo del Doctor Francisco José Mojica graduado en Ciencias Humanas Universidad de París V René Descartes (Sorbona) y la conferencia Posgrados Mentalidad de Innovación Exponencial, a cargo de la Doctora Edna Rocío Bravo Ibarra. Posteriormente, se efectuaron mesas de trabajos sobre los desafíos de la formación posgradual, por áreas de conocimiento. Los coordinadores de posgrado de la Universidad Francisco de Paula Santander Seccional Ocaña, participaron en el evento tanto de manera presencial como virtual.”

And its translation:

“The Colombian Graduate Network scheduled the First Meeting of Graduate Coordinators by area of knowledge, held on June 15, 2022 in hybrid mode (face-to-face and virtual) and aimed at graduate coordinators of the institutions that are part of the Colombian Graduate Network. The objective of the event was to share experiences and good practices in the academic and administrative management of postgraduate programs offered by the HEIs that are part of the Network, in different areas of knowledge. During the event, two conferences were held The "analogical" world of the past, facing the "intelligent" world of the future, by Dr. Francisco José Mojica, graduate in Human Sciences, University of Paris V René Descartes (Sorbonne) and the conference Postgraduate Exponential Innovation Mentality, by Dr. Edna Rocío Bravo Ibarra. Subsequently, round tables were held on the challenges of postgraduate education, by areas of knowledge. The graduate coordinators of the Universidad Francisco de Paula Santander Seccional Ocaña participated in the event both in person and virtually.”

3. RESULTS

The results of analyzing the documents of the 3 databases are shown from the general first to then analyze each of the sentences. Figure 2, presents the value of the emotions of each document in the official language of the UFPSO (Spanish) and the translation to English, where according to the figure the values of the emotions decrease as in 8 of the 9 documents with a percentage of almost 88.88% in the decrease of favorable emotions, with which it is clear that if it changes the final result of each document and the message may have a different context than it is in the Spanish language.

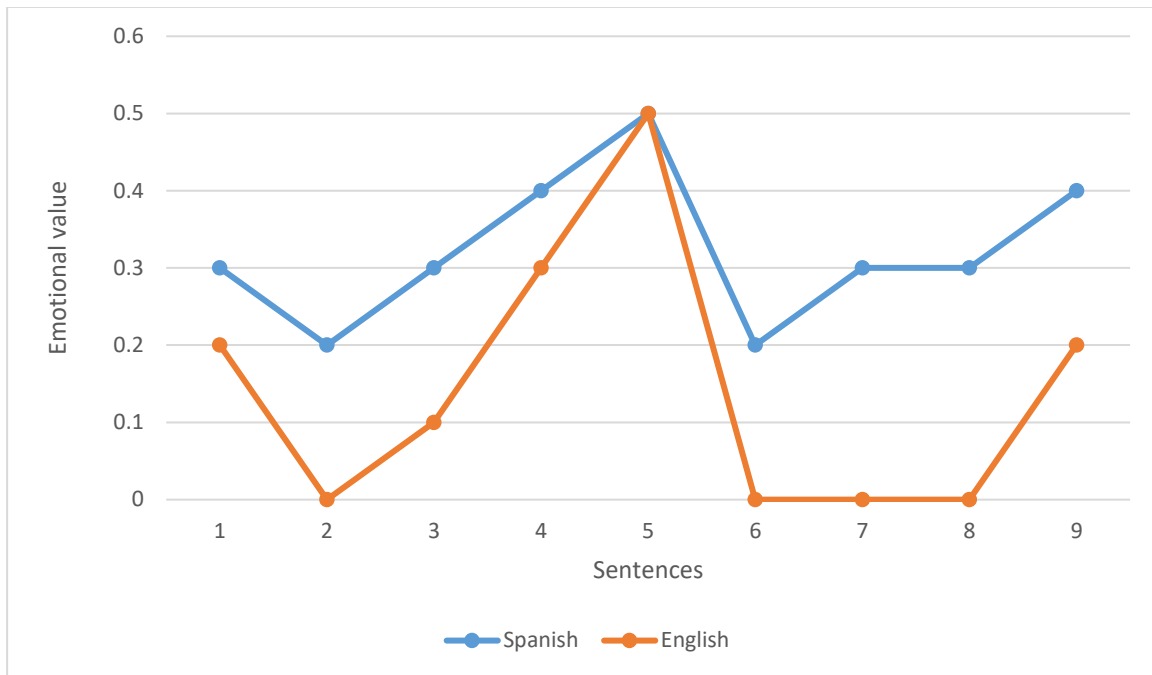
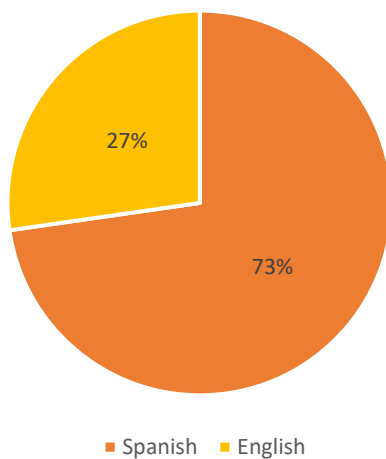
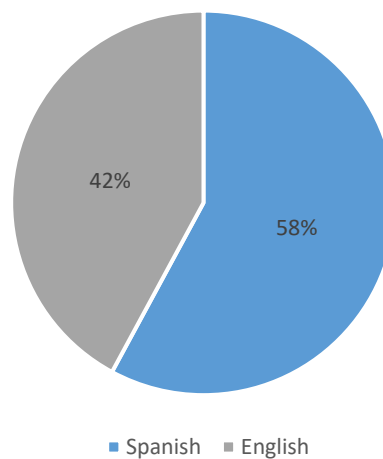


Figure 2. Value of emotions by language.

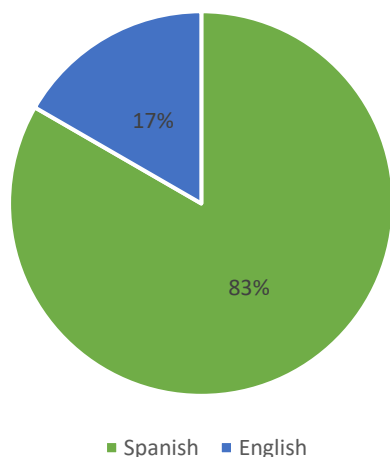
When analyzing the results of each database, as shown in Figure 3, it is determined that the first one framed within engineering areas and web news are the ones that have a very significant change by language with almost a 50% decrease, while the second database the level is more neutral, although always the difference is 16%, with which the terms of finance and economy are more generalized worldwide and are better interpreted regardless of the language.



(a)



(b)



(c)

Figure 3. Percentage of favorable emotions according to database: a) engineering, b) finance, c) news history.

A more detailed study of each document and the number of sentences can be seen in Table 1 and their value in the Spanish documents, and in Table 2 in English. The blue color shows the sentences that maintain their level of favorable emotion, while the yellow color shows those that increase in English, where it is clear that the decrease is evident when translating the documents into the non-official language.

Table 1. Value of emotions in Spanish.

Sentences	Document Spanish								
	1	2	3	4	5	6	7	8	9
1	0.5	0.2	0.3	0.5	0.5	0.2	0.2	0.2	0.5
2	0.1	0.2	0.2	0.3		0.3	0.6	0.4	0.3
3	0.6		0.3	0.5		0.1	0.2	0.3	0.4
4	0.1					0.3	0.3	0.1	0.1
5	0.3					0.1	0.4		0.3
6									0.5
7									0.8
8									0.5

Table 2. Value of emotions translation in English.

Sentences	Document English								
	1	2	3	4	5	6	7	8	9
1	0	0	0.2	0.5	0.5	0	0	0	0.1
2	0	0	0.1	0		0.1	0.1	0.1	0
3	0.7		0.1	0.3		-0.1	0	0.1	0
4	0					0	0	0	0.2
5	0.4					-0.2	0.1		0

6									0.2
7									0.6
8									0.7

4. CONCLUSIONS

The emotions of a document or opinion may vary according to the language, due to the lack of vocabulary with words that have a similar meaning when translated into a second language. Even the same document with the same language can vary due to its own dialects. The results show that when analyzing the most downloaded and read information documents at UFPSO written in Spanish and translating them into English, the level of emotions decreases. Although the value does not noticeably affect the message, the sentences and messages become more neutral in English, which may decrease the level of emotions.

References

- Agámez, S. M. G., & Gùzman, J. D. C. (2015). Influencia de los postulados de Taylor y Fayol en el desarrollo de las técnicas modernas de administración. *Revista Científica Profundidad Construyendo Futuro*, 3(3), Art. 3. <https://doi.org/10.22463/24221783.2259>
- Carbonell, V. C.-, Arias, Á. F. H.-, & Arias, A. C. S.-. (2020). Cobertura de las TIC en la educación básica rural y urbana en Colombia. *Revista Científica Profundidad Construyendo Futuro*, 13(13), Art. 13. <https://doi.org/10.22463/24221783.2578>
- Carrascal-Carrascal, H. (2011). El cubo Soma: Desarrollo del pensamiento lógico e intuición espacial. *Revista Ingenio*, 4(1), Art. 1. <https://doi.org/10.22463/2011642X.2172>
- Criado, M. Á. (2019, diciembre 19). Las emociones no significan lo mismo en todos los idiomas. *El País*. https://elpais.com/elpais/2019/12/19/ciencia/1576769294_589387.html
- Durán-Chinchilla, C. M., Cárdenas-García, M., & Velasquez-Pérez, T. (2016). Los modelos pedagógicos y su influencia en la práctica docente de la Universidad Francisco de Paula Santander. *Revista Ingenio*, 9(1), Art. 1. <https://doi.org/10.22463/2011642X.2068>
- Florez-Solano, E., Cardona-Duarte, N., & Clavijo-Ovalle, C. (2014). Diseño de un molino de martillos para la producción de pegamento cerámico. *Revista Ingenio*, 7(1), Art. 1. <https://doi.org/10.22463/2011642X.2030>
- Gutiérrez, M. F. O.-, Henao, A. M. G.-, Rodríguez, P. D. V.-, Peña, A. V. C.-, & Leguizamo, J. D. D.-. (2021). Teoría del desarrollo organizacional: Ineficiencia laboral y organizativa. *Revista Científica Profundidad Construyendo Futuro*, 14(14), Art. 14. <https://doi.org/10.22463/24221783.3162>
- Mavrou, I., & Saralegui, M. A. (2018). Percepciones sobre el significado emocional y las lenguas preferidas para tratar temas emocionales en hablantes bilingües y multilingües. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 12(25), Art. 25. <https://doi.org/10.26378/rnlael1225199>

- nuteco. (2019, julio 8). Traducciones que generan emociones. Empresa de Traducción | Online Traductores. <https://www.empresadetraduccion.es/traduccion-que-generan-emociones/>
- Orts, M. Á. (2020). Un estudio traductológico del sentimiento en los informes financieros español/inglés: Las emociones en la economía desde la perspectiva del Análisis del Sentimiento y la Teoría de la Valoración. *Sendeban*, 31, 395-417. <https://doi.org/10.30827/sendeban.v31i0.11817>
- UFPSOa. (2022). UFPSO - UFPS Seccional Ocaña presente en el Primer Encuentro de Coordinadores de Posgrados. https://ufpso.edu.co/new/UFPS-Seccional-Ocana-presente-en-el-Primer-Encuentro-de-Coordinadores-de-Posgrados_4485
- UFPSOb. (2022). UFPSO - Inscríbete en el curso de introducción a los sistemas de información geográfica con ArcGIS. https://ufpso.edu.co/new/Inscribete-en-el-curso-de-introduccion-a-los-sistemas-de-informacion-geografica-con-ArcGIS_4484
- UFPSOc, C. de M.-. (2022). UFPSO - Departamento de Derecho y Ciencias Políticas otorga reconocimiento a docentes por su excelencia. https://ufpso.edu.co/new/Departamento-de-Derecho-y-Ciencias-Policas-otorga-reconocimiento-a-docentes-por-su-excelencia_4486