

Jaime Sánchez Carnicer

Phraseology within the agri-food industry, *torreznos* and *adobados*: A corpus-based study

Abstract: The development and importance of the agri-food industry in recent decades has led to an increase in the number of exports of many products (MAPA, 2021), which has led many companies to consider the need to internationalize. Therefore, the offered resources should be of high quality so that the final recipients will identify them as their own. To obtain linguistic patterns between languages, such as Spanish and English, we have tools such as corpora that provide real examples of use (Seghiri, 2017, 2020; Sánchez Ramos, 2020, Ortego Antón, 2022). Consequently, in this work, through the analysis and exploitation of a comparable virtual corpus, C-MARMEAT, and a parallel virtual corpus P-MARMEAT (Ortego Antón, 2022), consisting of descriptive sheets in English and Spanish of products obtained from different companies, we aim to obtain an approximation of the phraseology related to *torreznos* and *adobados* in both languages. The results obtained from this analysis will allow us to obtain patterns of behaviour and to observe whether those existing in the translated texts are used in the originals written in English. With the results obtained, we will draw conclusions that will help translators and interpreters working in the agri-food industry identify and understand the patterns of behavior in English and apply them in their translation practice.

Keywords: Agri-food industry, corpus, phraseology, English, Spanish

Acknowledgements: This work has been carried out in the framework of the national R&D project entitled “Controlled natural languages, collaborative communication, and bilingual textual production in 3.0 environments” (PID2020-114064RB-I00), coordinated by Dr. Noelia Ramón García (University of León), and partially within the following projects: “TorreznosTRAD” (PROYEMER-2021-028), coordinated by Dr. María Teresa Ortego Antón and funded by the Emerging Projects Grants (Call 2021) of the University of Valladolid, “Multi-lingual and Multi-domain Adaptation for the Optimisation of the VIP system” (VIP II) (Ref. PID2020-112818GB-I00, Spanish Ministry of Science and Innovation), led by Prof. Gloria Corpas Pastor (University of Malaga) and “App para entrenar en posesición de traducción automática neuronal mediante la gamificación en entornos profesionales (GAMETRAPP) (TED2021-129789B-I00, Spanish Ministry of Science and Innovation), led by Prof. Cristina Toledo Baez (University of Malaga).

Jaime Sánchez Carnicer, CITTAC, University of Valladolid, e-mail: jaime.sanchez.carnicer@uva.es

1 Introduction: Agri-food industry and translation

The agri-food industry, also known as the agri-food industry, is one of the largest business industries, both at European level, where it accounts for 14.2% of manufacturing industry (FoodDrink Europe, 2021) and in Spain, since, according to data from the Spanish Ministerio de Agricultura, Pesca y Alimentación, MAPA (2022), this industry is the leading manufacturing branch of the industrial sector, representing 2.5% of the Gross Domestic Product (GDP) in Spain. All these data only endorse the idea of the importance of everything related to this industry in our society, as it is of a great weight of our economy, as well as to a high number of small and medium-sized enterprises, most of which have their main activity in this industry, which accounts for more than half a million workers according to MAPA (2022) data.

Similarly, the fact that this industry represents such a large economic and social volume is synonymous with its internationalization, i.e. the companies that produce goods in this industry do not limit themselves to attracting potential customers within the borders of the countries in which they operate, but also seek to be present in other markets. In the case of exports within this industry in Spain, more than 37 billion euros in exports were accounted for (MAPA, 2021).

Among them, as it can be perceived in the title, the subsector on which we will focus in this chapter is related to pork. According to MAPA data (2022), both in 2020 and 2021, its exports had a value of over 5 billion euros, occupying the first position both years. This shows the relevance of internationalization for companies in this subsector and the importance of projects such as TorreznoTRAD, which is the framework for the research described in this chapter and which, through the correct use of the language, boosts the number of recipients and clients.

However, the importance of the agri-food industry is not only reflected in our society by the economic aspect, but it has also generated interest in other fields, such as advertising (Rodríguez-Zúñiga and Soria, 1990; Díaz-Méndez and González-Álvarez, 2013; Carmona and Anguita, 2021) or linguistics, where we can find research focused on meat producers (Ortego Antón, 2019–2021), wine (Ramírez Almansa, 2019) or cheese products (Labrador and Ramón, 2015).

Despite all the research that has been carried out in recent years within this industry, there is a need for more, due to the evolution of the industry and an increasing demand for internationalization. Therefore, it is within this context that we would like to make the contribution presented in this chapter, limiting our study to the phraseology of *torreznos* and *adobados*.

Consequently, we will rely on the exploitation of the C-MARMEAT corpus, a comparable bilingual Spanish-English virtual corpus composed of product description sheets from different companies, and of P-MARMEAT, a parallel bilingual English-Spanish virtual corpus, consisting of texts belonging to the same textual genre.

We will briefly describe the project in which these corpora have been compiled, as well as the characteristics of each of them (Section 2), and then we will point out the methodology that will be used to extract and analyze the phraseology according to the parameters we will establish for this purpose (Section 3). In Section 4, we will show and describe the results obtained by applying this methodology and, finally, we will draw the conclusions (Section 5), as well as the references used (Section 6).

2 TorreznosTRAD and the C-MARMEAT y P-MARMEAT corpora

In recent decades, corpora have proved to be a very useful tool for translators and interpreters, as Corpas Pastor (2012, p.11) points out, as they allow the user to have access to a large number of units of meaning and specific functions in the register and textual form of the texts that contribute the corpus, both in the source language and in the target language. Due to the relevance that corpus linguistics has acquired within Translation Studies, with multiple researches highlighting its advantages (Ortego Antón, 2019–2021; Sánchez Carnicer, 2022; among others), in this paper we are going to extract data from C-MARMEAT, a comparable Spanish-English bilingual virtual corpus, and P-MARMEAT, a parallel English-Spanish bilingual virtual corpus, composed of product description sheets from different companies in the meat sector. These corpora, whose characteristics will be explained below, have been compiled within the TorreznosTRAD project, whose main objective is the development of a semi-automatic corpus-based application that assists in the writing and translation from Spanish into English of a specific textual genre: the description sheets of *torreznos* and *adobados*.

2.1 C-MARMEAT y P-MARMEAT: compilation and characteristics

The C-MARMEAT and P-MARMEAT corpora are composed, as shown in Table 1, of 100 texts in each language. In the case of the former, C-MARMEAT has a total of 37 860 cases or tokens¹ (if we break it down by language, English would have 24 462

¹ We understand case or token as “any instance of a particular wordform in a text” (McEnergy y Hardie, 2012: 50).

and Spanish 13 218, respectively). In the case of the other corpus, P-MARMEAT, it has a total of 29 528 cases or tokens (12 501 in English and 17 027 in Spanish).

Table 1: C-MARMEAT and P-MARMEAT corpus size.

CORPUS	Nº TEXTS EN	Nº CASES EN	Nº TEXTS ES	Nº CASES ES	TOTAL CASES
C-MARMEAT	100	24 642	100	13 218	37 860
P-MARMEAT	100	12 501	100	17 027	29 528

These two corpora from which we will extract the data have been compiled following the methodology proposed by Seghiri (2017) in which four phases are established: 1) Search and access to information; 2) Download; 3) Standardization and 4) Storage. In the following paragraphs, we will briefly explain the process followed in each of the phases to compile the two corpora:

- Search and access to information: Both corpora are made up of texts belonging to a specific textual genre, understood as “the internal structure and organization of texts from a specific variety” (Biber et al, 2007: 9), which is that of product fact sheets. We understand this genre as one that showcases a clear informative intention, with the possibility in some cases of also presenting a persuasive function. This genre showcases information about the product that can make the consumer opt for it (Ortego Antón, 2020). In the case of these two corpora, the texts come from the websites of meat companies, which were accessed through a keyword search on Internet search engines.
- Download: HTML files were downloaded from the different web pages and then TXT documents were created, with UTF 8 encoding, in order to use them in the analysis tools that we will use to extract the results.
- Standardization: Once all the texts were in, it was checked that they were all in the same format and with the same coding, and those that did not meet the established criteria were standardized.
- Storage: Folders and subfolders were created to store the texts, according to the corpus and the language in which the texts were written. Similarly, at the same time, they were coded with an alphanumeric code so that all the documents had the same name extension and were easily identifiable by their origin and the language in which they were written..

Furthermore, these two corpora have several characteristics, which are listed in Table 2. Some of these characteristics are shared, while others are particular to each one, since, although they are compiled following the same process, they are

not identical and the same process has not been followed with them once they have been assembled.

Table 2: Characteristics of C-MARMEAT y P-MARMEAT corpora.

C-MARMEAT	P-MARMEAT
	Virtual
Comparable	Pararell
	Bilingual
	Specialized
Tagged	Not Tagged
	Balanced and representative

In the case of the characteristics they share, we find the following: virtual, due to the fact that the texts of both corpora have been extracted from the websites of different companies, as well as their storage, which takes place in a digital environment; bilingual, both corpora are composed of texts in two languages, in this case, Spanish and English; and balanced and representative, as both meet the first of the characteristics, balance being understood as “proportions of data in our corpus reflect, in some way, the numbers of each type of interaction of interest that actually occur” (McEnery and Hardie, 2012: 8–9). In the case of representativeness, understood as “the minimum number of documents or words that a given corpus must contain in order to be considered valid and representative of the population to be represented” (Corpas Pastor and Seghiri, 2007: 166), both corpora are representative, having been analyzed using the ReCor² program (Arce Romeral and Peñuelas Gil, 2022).

If we look at the differences between the corpora, we can observe two: the first one would be in the texts that make up the corpora, since in the case of C-MARMEAT they are originally written in the language they are collected in for our corpus and they deal with the same subject, being a comparable corpus. In contrast, the P-MARMEAT corpus is a parallel corpus, i.e. the Spanish texts are originally written in Spanish, whereas the English texts are translations of them. The other difference lies in a step after the compilation of the corpus, in this case, the decision to tag the texts, as it has been done in C-MARMEAT, using the Open-Tagger program (Sanjurjo-González and Andaluz-Pinedo, 2021).

² Designed by Seghiri (2006).

Having described the size and detailed the characteristics of C-MARMEAT and P-MARMEAT, as well as the project for which they have been compiled, we proceed to detail the methodology that we will use in the analysis.

3 Methodology of analysis

Despite the usefulness of corpora as an essential tool in translation work, it is still necessary to have tools that allow access to the information and extract it as quickly and efficiently as possible in order to, in this case, proceed to analyze the sample of analysis, constituted by phraseology, defined by García Rodríguez (2019: 44) as the discipline that “studies the expressions formed by two or more words separated in writing, whose fundamental characteristics are pluriverbality, lexicalization, institutionalization, fixation and idiomaticity, the latter two to a certain degree”, of *torreznos* and *adobados*.

Therefore, in order to extract phraseological units, defined by Corpas Pastor (2003: 134) as the “stable combination of at least two words, which, by virtue of different currents will have as an upper limit the syntagm or the compound sentence and will present inherent features fixation and idiomaticity on their own, or a combination of both criteria”, which contain the two previously mentioned terms we will use as a starting point the methodology used by Sánchez Carnicer (2022), later replicated in other works (Ortego Antón and Sánchez Carnicer, 2023; Sánchez Carnicer and Peñuelas Gil, 2022, among others).

For this purpose, we will use the Sketch Engine tool, defined as:

[. . .] a corpus tool which takes as input a corpus of any language (with appropriate linguistic markup), and which then generates, amongst other things, words sketches for the words of that language. Those other things include a corpus-based thesaurus and ‘sketch differences’, which specify, for two semantically related words, what behaviour they share and how they differ (Kilgarriff et al. (2004: 105)

Within this tool, in which we have previously introduced both corpora, with the function N-GRAMS (Figure 1), understood as “sequences of elements as they appear in texts. These elements can be words, characters, POS tags, or any other elements as they encounter one after another in text” (Sidorov et al., 2014: 853), we will manually extract the phraseological units formed by between two and five most common words within the P-MARMEAT corpus containing one of these two terms or derivatives thereof, *torreznos* and *adobados* in Spanish, as well as their equivalent in English.

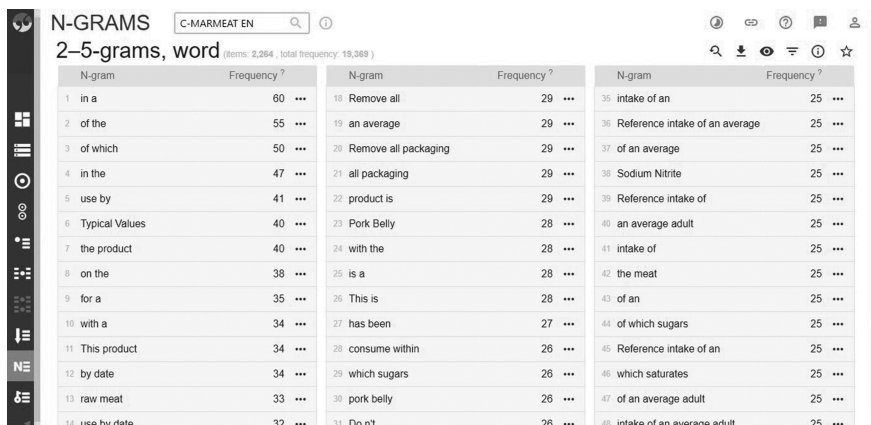


Figure 1: N-GRAMS in Sketch Engine.

Consequently, as shown in Tables 3 and 4, the analysis sample is made up of the following phraseological units:

Table 3: Analysis sample of *torrezno*.

ES	EN
torrezno de Soria	torrezno de Soria
auténtico torrezno de Soria	authentic soria torrezno
virutas de torrezno	torrezno de Soria chips

Table 4: Analysis sample of *adobado*.

ES	EN
adobado extra	marinated (pieza) extra
adobado con especias	marinated with spices
proceso de adobado	marinating process
proceso de adobo	seasoned with
adobar de madera tradicional	marinated in a traditional way
adobado ibérico	Iberian adobo
adobado tradicional de pueblo	homemade marinade

Once we have extracted our analysis sample, from which we will collect its occurrences, as well as calculate its normalized frequency, we will proceed to search for the English equivalents in the English language section of the other compiled

corpus, C-MARMEAT, also within the Sketche Engine tool to detect whether the phraseological units used in the translation of the product description sheets are also used in the texts of this textual genre when they are originally written in English.

4 The case of *torrezno y adobados*

As we have explained in the methodology section, we will first present the results obtained in the analysis of the parallel corpus, P-MARMEAT, in which we will search for N-GRAMS in Spanish and their equivalents in English, and then we will search for these equivalents in C-MARMEAT. Finally, we will compare the results gathered in both searches in a global way.

4.1 Phraseological units in P-MARMEAT

As far as the P-MARMEAT corpus is concerned, we have detected eight phraseological units (PUs), three in which the term *torrezno* appears and the other five composed of the term *adobado*, together with their equivalents in English. Table 5 shows their occurrences, as well as the normalized frequency of each of them.

Table 5: Phraseological units in P-MARMEAT.

P-MARMEAT (ES)			P-MARMEAT (TEN)		
UF	OCCURRENCES	FREQUENCY	UF	OCCURRENCES	FREQUENCY
torrezno de Soria	15	0.07	torrezno de Soria	14	0.086
auténtico torrezno de Soria	3	0.014	authentic soria torrezno	2	0.012
virtas de torrezno	2	0.009	torrezno de Soria chips	2	0.012
adobado extra	15	0.07	marinated (pieza) extra	2	0.012
adobado con especias	12	0.056	marinated with spices	2	0.012

Table 5 (continued)

P-MARMEAT (ES)			P-MARMEAT (TEN)		
UF	OCCURRENCES	FREQUENCY	UF	OCCURRENCES	FREQUENCY
proceso de adobado	9	0.042	marinating process	1	0.006
proceso de adobo	3	0.014	seasoned with	3	0.018
adobar de manera tradicional	3	0.014	marinated in a traditional way	3	0.018
adobado ibérico	3	0.014	Iberian adobo	2	0.012
adobado tradicional de pueblo	2	0.009	homemade marinate	4	0.024

As it can be seen in the table, in the case of the phraseological units containing the term *torrezno*, this has not been translated, leaving the term in Spanish, and the rest of the components of the unit have been translated. In the case of those which contain *adobado* or derivatives, we observe that there are some equivalents which maintain part of the unit in Spanish (*adobo ibérico* and its equivalent “Iberian adobo”) and, among those which present a complete translation, we observe two verbs in English: to marinate, the most common in the phraseological units in this language, and to season, present in only one unit.

Likewise, we also find differences in the equivalents of some of the terms that make up the unit, as is the case of *tradicional* in Spanish, in the units of *adobado de manera tradicional* and *adobado tradicional de pueblo*, which has two equivalents in English, “traditional way” and “homemade”. We also find differences in the units of *torrezno*, since in most of the equivalents, the preposition *de* is included between *torrezno* and *Soria*, while in the case of “authentic *soria torrezno*”, the term has not been kept the same as in Spanish, and the capital letter is not used in the proper noun.

If we look at the differences in the number of occurrences of the units in the two languages, we see that in the case of those formed by *torrezno*, they are practically identical in both languages. However, a greater difference is detected in those formed by *adobados*, since, although there are units with the same number of occurrences, there are others with very different numbers between languages, as in the case of *adobado extra*, *adobado con especias* and *proceso de adobado*,

which have a much higher number of occurrences in Spanish. Similarly, the opposite is the case, as the unit “homemade marinade” has a greater number of occurrences than its Spanish equivalent, *adobado tradicional de pueblo*, although, in this case, the difference between them is less than that between the other units.

4.2 Phraseological units in C-MARMEAT

Once we have extracted and analyzed the phraseological units present in both languages within the parallel P-MARMEAT corpus, we will search within the comparable corpus, C-MARMEAT, for the equivalents found in English, in order to observe whether they are used in texts originally written in that language. Table 6 shows the number of occurrences and the standardized frequency of each of them.

Table 6: Phraseological units C-MARMEAT.

C-MARMEAT		
UF	OCCURRENCES	FREQUENCY
torrezno de Soria	0	0
authentic soria torrezno	0	0
torrezno de Soria chips	0	0
marinated (pieza) extra	0	0
marinated with spices	1	0.003
marinating process	0	0
seasoned with	5	0.015
marinated in a traditional way	0	0
Iberian adobo	0	0
homemade marinade	0	0

In the case of the units formed by *torrezno*, in the C-MARMEAT corpus we did not find any occurrences, since there seem to be no texts on this product in the corpus; however, we did find examples of the use of *authentic* as an adjective used to define meat products, i.e. with the same function it fulfils in the phraseological unit found in P-MARMEAT. We also find occurrences of chips, although in this case it does not refer to an edible element, but with the meaning of “piece of wood”, so that its use would be different in both languages and corpus.

If we look at the units in which the term *adobado* appears, we observe that only two of them appear in the C-MARMEAT corpus, marinated with spices (1 occurrence, standardized frequency of 0.003), with a residual presence, and sea-

soned with (5 occurrences, standardized frequency of 0.015). As in the case of those formed by the other term in our study, we find examples in which part of the elements of the phraseological unit are included in the corpus, as in the case of *marinated together with spices*, with the difference that instead of the preposition with which we find in P-MARMEAT, the construction would be *marinated in a blend of spices*. Another example of this would be that of traditional way, which would appear with the verb to prepare instead of *marinate*, as is the case of the unit that we have detected as equivalent in P-MARMEAT. Finally, we also find a change of function in the use of one of the equivalents, namely *extra*, since in the units detected in P-MARMEAT in both languages it referred to a category of the product; however, in C-MARMEAT the word *extra* functions as an adjective to highlight some of the qualities of the product to which it refers.

4.3 Comparison of results

Once the phraseological units in English detected as equivalent in P-MARMEAT have been analyzed in C-MARMEAT, the results show that they are not used in the same way when the text is originally written in one language as when it is a translation from another language, in this case from English into Spanish. If we look at the composition of the units, the presence of adjectives stands out, which denotes the interest present in this textual genre in attracting the attention of the receiver to a certain extent and making the product more eye-catching.

On the other hand, when analyzing the phraseological units in the comparable corpus, we have detected that, although the unit does not appear as such in the corpus, many of its components do appear and, in many cases, with the same function as in the unit analyzed; nevertheless, it is also the case that they have a different grammatical function and even refer to a totally different reality.

It is also worth noting the fact that there are phraseological units in the P-MARMEAT which maintain some of their elements in the Spanish language without translation; however, this situation does not occur in the comparable corpus, since the terms are not maintained in another language.

To conclude this section, we would like to point out the richness of both languages to describe the product, that resides in the number of possible different phraseological units that we have collected in the small sample analyzed in this work, which implies a greater difficulty when transferring the information from one language to the other, making it necessary for translators and interpreters to be aware of them when carrying out their work.

5 Conclusions

The fact that corpora have become an essential tool in the world of translation, as well as having used a methodology based on the compilation and subsequent exploitation of a parallel bilingual Spanish-English virtual corpus, P-MARMEAT, and a comparable bilingual Spanish-English virtual corpus, C-MARMEAT, both composed of descriptive cards of meat products, has brought us the possibility of observing the behavior of the most common phraseological units of the terms *torrezno* and *adobado* in both languages through this work.

Likewise, the results obtained show us the differences between the two corpora, as we can appreciate that in the comparable corpus hardly any of the phraseological units detected in the parallel corpus appear, which clearly shows us the differences between the real use of the language, depending on the text, original or translated. Consequently, this makes us ponder on the veracity and usefulness of the parallel corpora, as in this case it shows the presence of phraseological units which are hardly used in the texts originally written in English. Similarly, we have also detected cases in which the grammatical function of some components of the units changes completely, as well as terms that acquire a totally different meaning depending on the corpus in which they are searched.

Therefore, we would like to make clear the greater usefulness of comparable corpora as opposed to parallel corpora, due to the possible errors or inadequacies of the latter, reflected both in the low presence of occurrences in C-MARMEAT and the use made of certain components of the phraseological units.

Finally, the results we have obtained lead us to the informed conviction that the tool proposed in the TorreznoTRAD project, where this research is framed. Companies need to be able to provide quality translations of their product description sheets that ensure that the phraseological units used as equivalents are commonly used and recognizable in the English language so that their products can be sold to a bigger number of consumers.

References

- Arce Romeral, L. & Peñuelas Gil, I. (April, 2022). *La compilación y explotación de C-MARMEAT: un recurso para la traducción y la comunicación multilingüe*. Paper presented in 39th International Conference of the Spanish Society for Applied Linguistics (AESLA), Las Palmas de Gran Canaria, Spain.
- Biber, D., Connor, U, & Upton, T. A. (Eds.) (2007). *Discourse on the move. Using corpus analysis to describe discourse structure*. Amsterdam: John Benjamins.

- Carmona Rodríguez, M. & Anguita Acero, J. M. (2021). El impacto de la publicidad en los hábitos alimenticios de los españoles: una cuestión de educación alimentaria. *Revista de comunicación y salud*, 11, 29–53.
- Corpas Pastor, G. & Seghiri, M. (2007). Determinación del umbral de representatividad de un corpus mediante el algoritmo N–Cor. *SEPLN: Revista de la Sociedad Española para el Procesamiento del lenguaje Natural*, 39, 165–172.
- Corpas Pastor, G. (2003). *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Iberoamericana.
- Corpas Pastor, G. (2012). Corpus, Tecnología y Traducción. In M. Casas & M. García Antuña (Eds.), *XII Jornadas de Lingüística* (pp. 75–98). Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- Díaz Méndez, C. & González Álvarez, M. (2013). La problematización de la alimentación: un recorrido sociológico por la publicidad alimentaria (1960–2010). *EMPIRIA: Revista de Metodología de Ciencias Sociales*, 25, 121–146.
- FoodDrink Eu (2021). *Data & Trends. EU Food and Drink Industry*. Brussels: FoodDrink EU.
- García Rodríguez, J. (2019). Las unidades fraseológicas del español y el catalán con elementos de la naturaleza: estudio cognitivo-contrastivo y propuesta de un diccionario electrónico de fraseología bilingüe. [Tesis Doctoral, Universitat Autònoma de Barcelona].
- Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine. *Proceedings of Eurlex*, 105–116.
- Labrador, B. & Ramón, N. (2015). Perfectly smooth, creamy and full flavoured: Online cheese descriptions. *Procedia-Social and Behavioural Sciences*, 198, 226–232.
- McEnery, T. & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Amsterdam/Philadelphia: John Benjamins.
- Ministerio de Agricultura, Pesca y Alimentación (2021). *Informe Anual de Comercio exterior agroalimentario y pesquero*. Madrid: Ministerio de Agricultura, Pesca y Alimentación.
- Ministerio de Agricultura, Pesca y Alimentación (2022). *Informe Anual de la industria alimentaria española. Periodo 2021–2022*. Madrid: Ministerio de Agricultura, Pesca y Alimentación.
- Ortego Antón, M. T. & Sánchez Carnicer, J. (2023). La terminología de la discapacidad de los sentidos y su traducción del español al inglés: un estudio basado en corpus. In Sánchez Ramos, M. M. (Ed.), *Investigaciones recientes sobre traducción y accesibilidad digital*. Berlin: Peter Lang.
- Ortego Antón, M. T. (2019). *La terminología del sector agroalimentario (español-inglés) en los estudios contrastivos y de traducción especializada basados en corpus: los embutidos*. Berlin: Peter Lang.
- Ortego Antón, M. T. (2020). Las fichas descriptivas de embutidos en español y en inglés: un análisis contrastivo de la estructura retórica basado en corpus. *Revista Signos. Estudios de Lingüística*, 53 (102), 170–194.
- Ortego Antón, M. T. (2021). E-DriMe. A Spanish-English frame-based dictionary about died meats. *Terminology*, 27(2), 294–321.
- Ortego Antón, M. T. (2022). *La investigación en tecnologías de la traducción. Parámetros de la digitalización presente y la posible incidencia en el perfil de los futuros profesionales de la comunicación interlingüística*. Berlin: Peter Lang.
- Ramírez Almansa, I. (2019). Vino, lengua y discurso. La interpretación de enlace (español-alemán-español) en el sector de la vitivinicultura. *Estudios franco-alemanes: revista internacional de traducción y filología*, 11, 151–164.
- Rodríguez Zúñiga, M. & Soria, R. (1990). La publicidad en el sector alimentario. *Revista de Estudios Agro-Sociales*, 154, 97–125.
- Sánchez Carnicer, J. & Peñuelas Gil, I. (April, 2022). *La traducción de «adobo» y otras técnicas de marinado al inglés: un estudio basado en corpus* Paper presented in 39th International Conference of the Spanish Society for Applied Linguistics (AESLA), Las Palmas de Gran Canaria, Spain.

- Sánchez Carnicer, J. (2022). *Traducción y discapacidad. Un estudio comparado de la terminología inglés-español en la prensa escrita*. Berlin: Peter Lang.
- Sánchez Ramos, M. M. (2020). *Documentación digital y léxico en la traducción e interpretación en los servicios públicos (TISP): fundamentos teóricos y prácticos*. Berlin: Peter Lang.
- Sanjurjo-González, H. & Andaluz Pinedo, O. (2021). OpenTagger: A flexible and user-friendly linguistic tagger. *Proceedings of the 56th Linguistics Colloquium*. Berlin: Peter Lang.
- Seghiri, M. (2006). *Compilación de un corpus trilingüe de seguros turísticos (español-inglés-italiano). Aspectos de evaluación, catalogación, diseño y representatividad [Tesis Doctoral, Universidad de Málaga]*.
- Seghiri, M. (2017). Metodología de elaboración de un glosario bilingüe y bidireccional (inglés-español/español-inglés) basado en corpus para la traducción de manuales de instrucciones de televisores. *Babel*, 63(1), 43–64.
- Seghiri, M. (Ed.) (2020). *La lingüística de corpus aplicada al desarrollo de la competencia tecnológica en los estudios de traducción e interpretación y la enseñanza de segundas lenguas*. Berlin: Peter Lang.
- Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A., & Chanona-Hernández, L. (2014). Syntactic N-grams as machine learning features for natural language processing. *Expert Systems with Applications*, 41(3), 853–860.