

María José Marín Pérez*

Department of English Philology
University of Murcia, Spain
mariajose.marin1@um.es

EXPLORING THE UNEXPECTED IN LEGAL DISCOURSE: A CORPUS-BASED CONTRASTIVE ANALYSIS OF SPANISH AND BRITISH JUDGMENTS ON IMMIGRATION

Abstract

In spite of being popularly regarded as examples of objectivity, two collections of Spanish and British judicial decisions related to the search terms *migration*, *immigration* and their Spanish equivalents were examined in search for evidence of the use of evaluative vocabulary, which appears to be considerably significant judging by the amount of such lexical items found in both corpora. This research thus introduces a contrastive corpus-based study of two legal corpora through the replication of the *appraisal* theory model. The frequency lists from both corpora, obtained using the software *Lancsbox* (Brezina et al., 2015) were compared by examining and classifying those vocabulary items amongst the top 2,500 types in the lists using the taxonomy provided by appraisal theory. The findings show that the British dataset contains a greater proportion of evaluative vocabulary, particularly as regards the category *affect* within the *appraisal* system. Such findings could be related to the very nature of its legal system, where the law is said to be judge-made, leaving greater freedom for the expression of stance as opposed to the Spanish system, which is codified and may somehow constrain legal actors in the way in which they convey their attitude towards the propositional content of legal texts.

148

Key words

legal language, corpus-based discourse analysis, appraisal theory, immigration, judicial decisions.

* Corresponding address: María José Marín Pérez, Universidad de Murcia, Departamento de Filología Inglesa, Campus de La Merced, Calle Santo Cristo s/n, 30071 Murcia/Spain.

1. INTRODUCTION

According to Eurostat,¹ the statistical office of the European Union, there were 34.3 million people born outside the EU living in European countries on 1 January 2015. The largest numbers could be found in Germany (7.5 million people), the United Kingdom (5.4 million), Italy (5 million), Spain (4.5 million), and France (4.4 million), 76% of the total amount of non-EU nationals in the whole of the EU.

Regrettably, non-EU immigration has recently become relevant owing to the humanitarian crisis of Syrian refugees, fleeing from the horrors of civil war. However, the response of the European authorities to this and other crises has been more than questionable. In Spain, less than 2,000 Syrian refugees have been granted asylum to date and the African borders are being invigilated by applying strongly restrictive policies which often go against the basic principles of human rights.

Yet, how do legal systems react to this or other migration flows into the EU? Could this be observed and quantified in legal text? These and other questions have motivated this research, whose major aim is to find linguistic evidence of the expression of appraisal in legal discourse through a contrastive analysis of two legal corpora.

It is commonly assumed that legal texts such as judicial decisions could be considered examples of objectivity, being always based on fact and leaving little room for the expression of attitudinal stance. However, after examining two corpora of British and Spanish judicial decisions related to the phenomenon of immigration, this does not seem to be the case or, at least, there seems to be enough evidence to consider that the expression of appraisal in this legal genre is more than mere chance, particularly as regards the British corpus.

Therefore, a hypothesis was formulated in order to try and account for the differences that might be found between the Spanish and the British corpora employed in this study. This hypothesis relates to the very nature of both legal systems. The Spanish legal system belongs in the realm of civil law – which is mostly codified – while the British one revolves around the principle of binding precedent, that is, it belongs in the common law system. In civil law systems, which derive from Roman law, judges cite the articles in the codes and apply them, which are less prone to interpretation. Conversely, in case law systems, judges follow the principle of binding precedent (or *stare decisis*) whereby they are bound by existing decisions or rulings that have set precedent. The law is said to be judge-made, to be interpreted by judges, which may allow for greater subjectivity in deploying legal discourse. As a consequence, British judicial decisions might display a greater number of linguistic items conveying attitude or appraisal than their Spanish counterparts.

In order to confirm this hypothesis, two research questions were formulated as follows:

¹ <http://ec.europa.eu/eurostat>

RQ 1: To what extent can legal text reflect attitudinal stance? How can this be quantified?

RQ 2: Do differing legal systems condition the way in which attitude is expressed by legal actors?

This research thus presents a corpus-based comparative study of two corpora of judicial decisions – of 600 texts each – issued by British and Spanish courts, which deal with the phenomenon of immigration. Its major aim is to try and find linguistic evidence of the way in which propositional information is presented within this legal genre through the application of the *appraisal* framework as defined by White (1999), Martin (2003), Eggins and Slade (1997), Rothery and Stenglin (2000), and Kaltenbacher (2006), and to relate the results to the hypothesis formulated above.

The second section of this study offers a literature review on corpus-based discourse analysis and its relation to legal discourse, and the third one introduces the methodology implemented to obtain the results, which are displayed and discussed in section four, followed by the conclusion in section five.

2. CORPUS-BASED DISCOURSE ANALYSIS AND LEGAL CORPORA

The description of legal language has traditionally been accomplished by reputed specialists (Alcaraz Varó, 1994; Borja, 2000; Mellinkoff, 1963; Tiersma, 1999) whose conclusions, deriving from deep knowledge and extensive expertise in the field, are usually based on intuition or on relatively reduced samples of texts. These seminal studies often present a top-down characterisation of the major traits of *legalese*, following a deductive approach whereby the rule usually precedes the actual description of the examples provided. In the recent years, there has been a growing tendency towards corpus-based and corpus-driven² descriptions of this English for Specific Purposes (ESP) variety, which provide a bottom-up characterisation of legal English (Breeze, 2015; Biel & Engberg, 2013; Marín Pérez & Rea Rizzo, 2012; Pontrandolfo & Goźdź-Roszkowski, 2014). Yet, there is still a clear need for a greater number of such studies.

Regarding the phenomenon of immigration, it has been examined as seen in public discourse (mostly in the media) with the aid of corpus linguistics techniques (Baker, Gabrielatos, & McEnery, 2013; Blinder & Allen, 2016; Gabrielatos & Baker, 2008, to name but a few). However, only Pérez-Paredes, Aguado Jiménez, and Sánchez Hernández (2017) have explored legal texts (UK legislation and official information) to scrutinise how immigrants are depicted within this domain. On the

² In corpus-based linguistic studies a query is formulated in advance so as to find evidence in a corpus, whereas corpus-driven analyses base their conclusions solely on linguistic findings obtained from corpora and adopt an inductive approach to language description.

other hand, none of the authors above have compared differing legal systems and, to the best of my knowledge, almost no contrastive corpus-based research has been carried out to date about this topic.

In addition, although some scholars have recently examined legal texts in search for evidence on the deployment of evaluative meanings (Finnegan, 2010; Goźdź-Roszkowski, 2017; Mazzi, 2010), only a fairly limited number of comparative studies have been pursued using corpus linguistics to account for the different mechanisms to express stance as displayed in legal text. Portrandolfo and Goźdź-Roszkowski (2014) and, more recently, Orts (2018) acknowledge the relevance of this feature of language and the differences across legal systems (American v. Italian, and British v. Spanish, respectively). This research is therefore intended to try and bridge this gap.

3. METHODOLOGY

3.1. Corpus design

As shown in Table 1, two corpora were employed in this study which comprise 600 judicial decisions each. However, the size of both text collections differs considerably as the Spanish corpus has 2.4 million tokens (or running words) while the British one is noticeably bigger, containing 3.7 million. As a consequence, the number of types, that is, the number of different words in a corpus regardless of the number of times they occur within it, is also higher, being 20,236 in the Spanish corpus and 25,268 in the British one. This is why, when presenting the results and analysing them, the figures will be normalised³ so as to allow for comparison, as indicated in section 4.

Two sources were employed to obtain the texts, which were produced between 2016 and 2017: the *CENDOJ*⁴ (Spanish Legal Documentation Centre) and the *BAILII*⁵ (British and Irish Legal Information Institute). Both databases offer free access to public judgments from Spanish and British courts.

CORPORA	TEXTS	TOKENS	TYPES
Spanish corpus (<i>CENDOJ</i>)	600	2,396,985 (2.4m)	20,236
British corpus (<i>BAILII</i>)	600	3,723,587 (3.7m)	25,268

Table 1. Corpora description

³ For the data to be normalised, the raw/absolute frequency of a type was divided by the number of tokens in the corpus and multiplied by 1,000,000. This simple operation makes two datasets of different size comparable.

⁴ <http://www.poderjudicial.es/search/indexAN.jsp>

⁵ <http://www.bailii.org/>

The selection of the texts in each corpus was carried out using the search engines included in the *CENDOJ* and the *BAILII* websites, which allow for advanced searches. The search terms used to obtain the Spanish texts included *inmigración* (immigration), *inmigrante* (immigrant), *extranjero* (literally *foreigner*) and *extranjería* (this term is employed in Spanish to refer to the laws and regulations on immigration, for instance, the *Ley de Extranjería 4/2000* [Immigration Act 4/2000]). In English, the search terms were *immigration*, *migration*, *immigrant*, and *migrant*. The corpora were processed using *Lancsbox* (Brezina, McEnery, & Wattam, 2015) to obtain the type lists and *TagAnt* (Anthony, 2015) to add POS (Part of Speech) tags to the words in each text collection.

3.2. Appraisal theory: The systematisation of evaluative language

As stated in the introduction, the analysis of both text collections necessarily required a clear systematisation of the information extracted from them. Such systematisation could contribute to reducing the degree of subjectivity implied in the qualitative interpretation of the data, which was automatically retrieved by the software in its entirety and validated both automatically and manually, as shown below. This necessity became particularly evident when attempting to define a clear procedure for the classification of vocabulary items that could potentially be employed to express attitudinal stance.

In that vein, appraisal theory provides a wide array of categories or systems that could be employed to that end. Following Martin (2003: 145), “the term *appraisal* will be used here for the semantic resources used to negotiate emotion, judgments and valuations alongside with resources for amplifying and engaging with these evaluations”. Martin’s definition covers both overt and implied appraisal, specifically, explicit and implicit mechanisms that could either express evaluative attitudes directly or rather evoke them. In this case, and owing to the fact that this is an automatic analysis of 1,200 texts, only the first option was considered.

The appraisal system is part of a linguistic model of analysis, systemic linguistics, whereby language is envisaged as a whole network of systems which are interconnected and available for the speaker to choose amongst them so as to express meaning. According to Halliday (2003: 8), language is conceived as a set of options which “arise from the environment of other options” thus constituting the language paradigm, in line with Saussure’s definition of linguistic relations. Those options, often identified as language functions, are defined as systems, where categories add to each other forming a complex network of meanings. One of those systems is *appraisal*.

This linguistic model could therefore be implemented to scrutinise the way in which information is presented on the part of the writer/speaker in legal text. It comprises four major categories or systems as established by White (1999), Martin

(2003), Eggins and Slade (1997), Rothery and Stenglin (2000), and Kaltenbacher (2006), namely, *affect*, *appreciation*, *judgment*, and *amplification*.

The category *affect* refers to those linguistic items which appeal or make reference to emotion (Martin & White, 2005). After processing the two corpora and obtaining the two type lists, as will be explained in greater detail below, different examples were included within this group (see Table 2) such as *suplicar* (beg), *preocupado* (worried), *vulnerable*, or *fearful*. When using the term *appreciation*, Eggins and Slade (1997: 125) state that it marks “the speakers’ reactions to and evaluations of reality”. Words like *arbitrario* (arbitrary), *controvertido* (controversial), *nice*, or *helpful* fall within this class. The *judgment* system comprises those terms which “present a speaker’s assessment of other people’s ethics, morality, or social values” (Kaltenbacher, 2006: 272), for instance, *indebido* (wrongful), *legítimo* (legitimate), *discriminatorio*, or *unfair*.

	SPANISH	ENGLISH
AFFECT Dis/inclination Un/happiness In/security Dis/satisfaction	<i>Exigir</i> (demand), <i>suplicar</i> (beg) <i>Feliz</i> (happy), <i>obligado</i> (obliged) <i>Preocupado</i> (worried), <i>ansioso</i> (anxious) <i>Harto</i> (fed-up)	<i>Expectation, opportunity, Free, forced, satisfied</i> <i>Vulnerable, emotional, fearful</i> <i>Like, hate</i>
APPRECIATION Reaction Valuation Composition	- <i>Arbitrario</i> (arbitrary), <i>controvertido</i> (controversial) <i>Preciso</i> (accurate), <i>Correctamente</i> (correctly)	<i>Nice, adequate, harsh</i> <i>Helpful, difficult, essential</i> <i>Detailed, precise, complex</i>
JUDGMENT Social sanction Social Esteem	<i>Indebido</i> (wrongful), <i>legítimo</i> (legitimate) <i>Constante</i> (tenacious), <i>inteligente</i> (intelligent)	<i>Discriminatory, unfair, true</i> <i>Determined, capable</i>
AMPLIFICATION Enrichment Augmenting Mitigation	- <i>Mucho</i> (much), <i>más</i> (more), <i>muy</i> (very) <i>Reducido, simplemente</i>	- <i>Entirely, much, highly</i> <i>Sole, shortly, mere</i>

Table 2. The appraisal theory framework: examples from both corpora

As illustrated by Table 2, the four main categories or systems are also divided into other subcategories or subsystems for a closer examination of the items included in each of them. The hyphens indicate that no words were found within that subcategory in either corpus.

3.3. Research methodology

The two corpora were processed with *Lancsbox* (Brezina et al., 2015) with the aim of obtaining the list of their most frequent 2,500 word types. The inventory of evaluative vocabulary which Kaltenbacher (2006) extracts and classifies according to the appraisal system from a corpus of tourism websites was used as reference. These items were automatically searched in both text collections. However, and due to the differing fields this study and Kaltenbacher's pertain to, a manual search was also required so as to spot other items which could encode interpersonal meanings specifically within the legal field.

Thus, the list of the most frequent 2,500 English and Spanish types was searched both automatically and manually, and a total amount of 420 word types (146 in Spanish and 274 in English) were selected for analysis once the list was filtered. This pre-selection phase implied not only reading the whole 2,500 type lists and selecting those that may potentially express appraisal, but also examining the context of usage of those words that might be ambiguous or polysemic. For instance, a word like the Spanish *impuesto* could refer to the English term *tax* or to the adjective *imposed*, only the latter interpretation could be considered evaluative. Therefore, only the frequency counts associated to that meaning were taken into account.

Filtering became necessary after the initial scrutiny of the lists at hand. As a matter of fact, authors (Baker et al., 2013; Gabrielatos & Baker, 2008; McEnery & Baker, 2016) frequently agree that the level of subjectivity involved in discourse analysis is often high, hence the need to keep it under control. The mere selection of the lexical items to be studied or classified as evaluative, as is the case, could vary considerably from one researcher to another. Moreover, Hunston (2006: 3) criticises how systemic linguistics may appear to constrain language analysis by "shoehorning" terms and expression into pigeonholes or categories which, on the other hand, also enrich language description by systematising it in an efficient way.

Following from the above, an attempt to objectivise the initial list of terms was made by designing an inter-rater reliability scale for two specialists to group the vocabulary inventory pre-selected for this research. They were requested to classify such terms into three categories, namely, *evaluative*, *possibly evaluative*, and *non-evaluative*. Only the overlapping items identified by both specialists for the first two groups were kept, the rest of of the items were discarded. The overlap percentage between the first two groups was 73.2% for the *evaluative* category and 54.1% for the *possibly evaluative*.

4. RESULTS AND DISCUSSION

4.1. Preliminary results

Having filtered the pre-selected elements by asking two specialists to examine and classify them, a significant difference could be observed between the Spanish and the English corpora, as displayed in Table 3. Firstly, the British texts contain more vocabulary items expressing appraisal than the Spanish ones, that is, 274 against 146. The proportion in both cases was calculated by using the whole type lists as reference, given that the two datasets differ in size and, as a result, they should not be compared in absolute terms. The items identified as evaluative were divided by the total amount of different words extracted by the software from each text collection, which was 20,236 in the Spanish set and 25,268 in the British one (see Table 1). The resulting percentage of evaluative vocabulary items was 0.72% for the former and 1.08% for the latter – 33% higher than the Spanish list – as shown in Table 3. Therefore, in a first approach to the data, our hypothesis seems to be confirmed.

Apart from obtaining the type list from a corpus, *Lancsbox* (Brezina et al., 2015), the software employed to process both corpora, can measure other parameters which may contribute to the portrayal of the lexicon under examination. As regards frequency, having normalised the figures for comparison, the items in the Spanish list, although less numerous, occur slightly more often than the English ones, 123.17 times against 116.81 respectively on average.

	# ITEMS	AV. FREQ.	EXAMPLES
SPANISH CORPUS	146 (0.72% of type list)	123.17	<i>Significativo</i> (significant), <i>preocupado</i> (worried), <i>grave</i> (serious), <i>negative</i> (negative), <i>riesgo</i> (risk), <i>controvertido</i> (controversial)
BRITISH CORPUS	274 (1.08% of type list)	116.81	<i>Error</i> , <i>appropriate</i> , <i>reasonable</i> , <i>failure</i> , <i>vulnerable</i> , <i>adequately</i>

Table 3. Preliminary results

Text distribution can also be calculated by processing the texts. Once the software has identified all the different words in each corpus (the types) and indicated how many times they repeat, it calculates the percentage of texts where each word can be found. This parameter, text distribution, may also inform of the relevance of a given term or terms within a text collection. The higher the percentage of texts a word occurs in, the more representative of a corpus it could be deemed. In this particular case and on average (there is greater variation when looking at the

words in isolation), the figures are almost identical, being 24.5%⁶ in the English corpus and 24.34% in the Spanish one. Therefore, it could be argued that, in spite of being more numerous, in general, vocabulary items showing appraisal could be considered similarly relevant in both corpora as they distribute likewise.

Still within the preliminary results section, we could continue to describe the linguistic behaviour of evaluative terms by examining the parts of speech (POS) or grammatical categories they belong in. To that end, both corpora were processed with *TagAnt* (Anthony, 2015), a POS tagger⁷, which recognises such grammatical categories almost on the fly. As illustrated in Table 4, there is a coincidence in the higher percentage of adjectives found in both corpora, ranking first out of four and six POS classes in English and Spanish respectively. Words such as *appropriate*, *reasonable*, *accesorio* (accessory), or *impuesto* (imposed) fall within this group. The second most numerous category in the English corpus are nouns like *challenge* or *refusal*, followed by adverbs such as *clearly* or *simply*. However, adverbs like *ciertamente* (certainly) or *relativamente* (relatively) occur more often in the Spanish corpus than nouns, although the difference is not significant. Verbs are the fourth grammatical category in both text collections, *reject*, *fail*, *despreciar* (despise) or *desvirtuar* (distort) instantiate their usage. Finally, we could only find examples of quantifiers and pronouns that fit within the *amplification* subcategory (see Table 2) in the Spanish corpus, although they represent a very small proportion.

	ADJS	NOUNS	ADVS	VERBS	QUANTS	PRONS
SPANISH CORPUS	54.8% <i>Absoluto</i> (Absolute) <i>Accesorio</i> (Accessory) <i>Impuesto</i> (Imposed)	16.44% <i>Subjetividad</i> (Subjectivity) <i>Arbitrariedad</i> (Arbitrariness) <i>Contradicciones</i> (Contradictions)	17.91% <i>Ciertamente</i> (Certainly) <i>Sólo</i> (Solely) <i>Relativamente</i> (Relatively)	8.31% <i>Debe</i> (Must) <i>Despreciar</i> (Despise) <i>Desvirtuar</i> (Distort)	1.27% <i>Pocas</i> (Few) <i>Muchas</i> (Many)	1.27% <i>Nada</i> (Nothing) <i>Ninguno/a</i> (None)
BRITISH CORPUS	41.3% <i>Appropriate</i> <i>Reasonable</i> <i>Likely</i>	23.81% <i>Challenge</i> <i>Refusal</i> <i>Error</i>	23.43% <i>Clearly</i> <i>Simply</i> <i>Entirely</i>	11.73% <i>Reject</i> <i>Fail</i> <i>Satisfy</i>	-	-

Table 4. POS category distribution

In sum and on the whole, evaluative vocabulary items tend to be more numerous in English judgments than in Spanish ones. Judging by the figures, a noticeably higher amount of lexical items expressing appraisal (33% higher) has

⁶ This percentage refers to the total number of texts in each corpus, which is 600 in both cases (see Table 1).

⁷ Freely available at <http://www.laurenceanthony.net/software/tagant/>

been found in a corpus of judicial decisions issued by British courts when compared with a similar collection of these texts obtained from Spanish sources. However, the average frequency of these items is higher in the Spanish dataset, though not significantly. As regards text distribution, which may point at the greater representativeness of a given word or words in a corpus, evaluative vocabulary items distribute similarly in both text collections, appearing, on average, in approximately 24% of the texts.

4.2. Results by categories

Figure 1 represents the proportion of evaluative vocabulary items in both corpora. Yet, what do all these figures tell us about the way in which legal argument is built as regards the expression of attitude in both systems? Comparatively and on the whole, as already stated, the British corpus displays a higher percentage of vocabulary items expressing appraisal: 1.08% against 0.72% in the Spanish corpus. Going back to our initial hypothesis, this could be explained by the greater freedom of British judges to interpret the existing precedent. The law is judge-made in case law systems, broadly speaking. Conversely, Spanish judges are rather limited by codified law in this respect, having to cite the articles contained in the codes, which may account for the lower number of these items in this text collection.

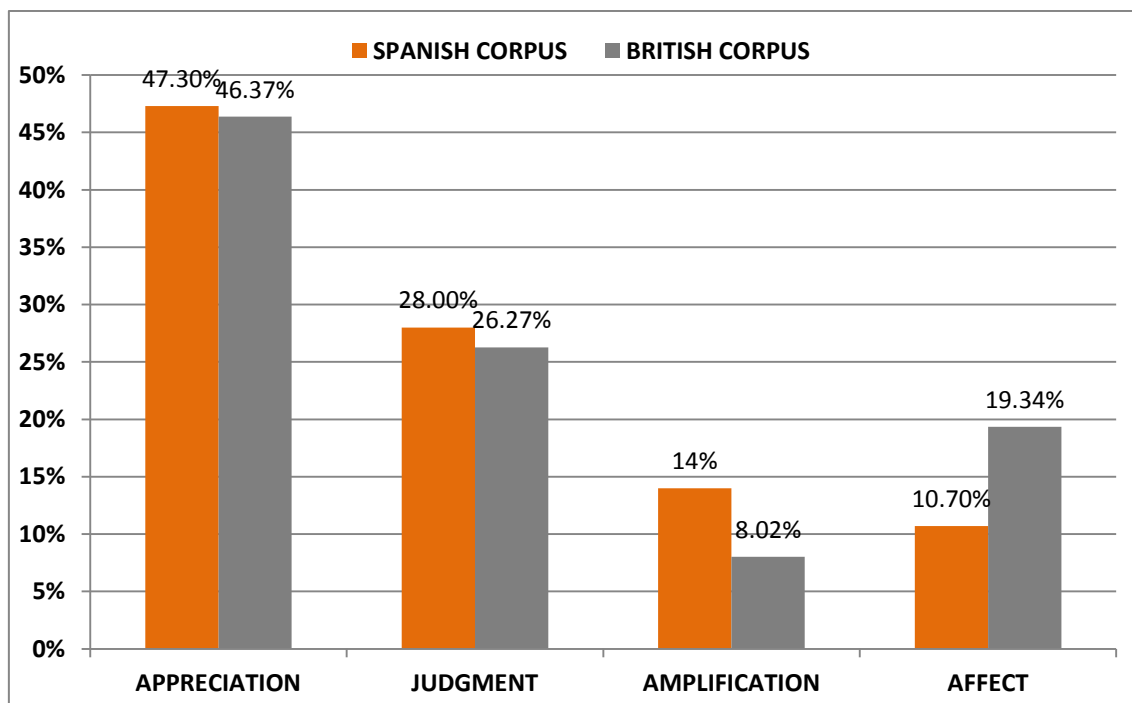


Figure 1. Proportional distribution of appraisal terms per category, expressed in percentages

Nevertheless, as illustrated by Figure 1, the proportion of vocabulary items expressing appraisal is similar in both corpora if we examine the categories separately, *appreciation* being the one which presents the highest number of instances, followed by *judgment* and *affect* (considerably more numerous in English) and finally, *amplification*. All the percentages shown in Figure 1 were calculated with respect to the total number of vocabulary items identified as evaluative in each corpus. If we compare them with the total amount of types/words in both corpora, they are more numerous in the British dataset in all categories and also on average. The analysis presented below will concentrate on the first three, namely, *appreciation*, *judgment*, and *affect*.

4.2.1. Appreciation

Appreciation ranks first with respect to the whole list of appraisal terms extracted from both text collections. It displays 47.3% of these items in the Spanish corpus and 46.37% in the British one. *Arbitrario* (arbitrary), *necesario* (necessary), *desprecio* (contempt), *controvertido* (controversial), or *pertinente* (relevant) are examples within this group in Spanish. If we examine their context of usage, we can appreciate how the speaker is showing attitudinal stance towards the content of the text:

1. (...) los acusados realizaron diferentes operaciones con la finalidad de usar **arbitrariamente** los fondos públicos que tenían a su disposición. Este uso **arbitrario** se hizo con **desprecio** total de las normas administrativas y legales. [(...) the defendants performed several monetary transactions with the aim of using the public funds at their disposal **arbitrarily**. Such **arbitrary** actions showed their utter **contempt** towards administrative and legal regulations (...)]
2. (...) con la finalidad de que las partes puedan hacer de ellas el uso procesal que pretendan en relación a su derecho de defensa, y que resulte **pertinente** a juicio del Tribunal. [(...) so that the parties can use such sources during the proceedings as regards their right of defence and as long as this fact is deemed **relevant** by this court (...)]

Something similar happens with *accurate*, *helpful*, *adverse*, *difficult*, or *critical*, which were found amongst the most frequent items in the English corpus within the same category:

3. *It may be necessary to grant stay if the impending appellate decision is likely to have a **critical** impact on the current litigation.*
4. *He made additional **adverse** findings. A medical report did not provide support for the account. He concluded (at [52]) that the account was 'entirely lacking in credibility'.*

5. *Relating to the Status of Refugees 1951 (United Nations) Art.1A(2). The House of Lords held that the Guidelines on International Protection, (...), provided a very **accurate** and **helpful** distillation of the effect of the relevant international authorities.*

Concerning the analysis of the data above and according to Eiggins and Slade (1997: 125), *appreciation* refers to “the speakers’ reactions to and evaluations of reality”. Consequently, it is not surprising to find a similar proportion of these elements in both corpora as, in my view, it is the judge’s role to evaluate and thus justify a decision as well as other legal actors’ to present their position/reaction to, for instance, evidence or witnesses’ statements. Due to this fact, it also appears reasonable that this category stands out as the most numerous one out of the four analysed in this study.

4.2.2. Judgment

Though not so numerous as the *appreciation* category, the expression of *judgment* (both as social sanction and social esteem) ranks second, being evenly distributed in both corpora (26.27% of the terms in the British corpus, and 28% in the Spanish one). Similarly to the *appreciation* system above, the proportion of these lexical items is comparable in both corpora. The Spanish words *exceso* (excess), *indebido* (wrongful), *indefensión* (lack of defence), or *gravedad* (seriousness) fall within this class. When examined in context, their judgmental load becomes even more evident:

6. *La consistencia de los indicios necesarios (...) no pueden hacerse depender directamente de la **gravedad** del delito que se pretende investigar, (...)* [The consistency of the necessary evidence (...) cannot depend directly on the **seriousness** of the offence which needs to be investigated (...)]
7. *Por lo tanto, nos encontramos con una clara decisión (...) que impide al interesado poder combatirla (...). Esta clara **indefensión** ha de llevar a la anulación (...)* [Therefore, we are faced with a clear decision (...) which prevents the defendant from combating it (...). Such a clear **lack of defence** (...)]

Unethical, wrong, inadequate, or disproportionate exemplify this category in the English corpus. Let us examine their contexts of usage:

8. *He further submitted that the district judge was **wrong** in failing to “grasp the nettle” but holding, as she did, that she should not proceed to determine the matter based on **inadequate** evidence.*
9. *An appellant will frequently not be represented at all given the restrictions on legal aid and the fact that many appeals will be conducted when the appellant is outside*

*the country. It is **unethical** for a legal representative to act on the basis of information he or she cannot discuss with the client.*

As Kaltenbacher puts it (2006: 272), the *judgment* system (or category) “presents a speaker’s assessment of other people’s ethics, morality, or social values”. This fact could thus explain that the category *judgment* ranks second, including one fourth of the items in the whole vocabulary list in both languages. In fact, this appraisal subsystem points at the speakers’ valuation on the content of the text, that is, on the facts that are being tried at court, hence its vital character in building legal discourse within this genre. Both judges and other legal actors like barristers, legal experts or even witnesses often have to determine the moral value of the acts being tried, and, in doing so, they may let judgments act as an ethical reference and an example to follow.

4.2.3. *Affect*

In spite of ranking third, this category displays the greatest differences between the two text collections, finding almost twice as many of these items in the British corpus (19.34%) as in the Spanish one (10.70%). The system *affect* includes Spanish terms which appeal or make reference to emotion such as *exigir* (demand), *suplicar* (beg), *feliz* (happy), *preocupado* (worried), or *ansioso* (anxious). They could be found in different contexts which illustrate some of the topics they are associated with:

10. (...) *la demandada (...) es de origen árabe (...) de nacionalidad marroquí, vivía en Agadir junto con sus padres, trabajaba como comerciante junto con su padre, tenía una vida tranquila y **feliz**.* [The defendant (...) has Arab origins, is from Morocco, and lived in Agadir with both her parents, working as a trader with her father, she led a quiet and **happy** life].
11. *El Abogado del Estado (...) se opuso a la demanda, (...) terminando por **suplicar** que se dictara Sentencia que desestime el recurso (...)* [The state advocate (...) contested the claim (...) ending up by **begging** for the court’s decision to dismiss the appeal (...)]

Amongst the most frequent lexical items on the English list we find *vulnerable*, *hate*, or *fearful*. These are some instances of their contexts of usage:

12. *In practice, **vulnerable** and marginalised groups do not have access to these documents and/or cannot get to registration centres. They therefore have difficulty obtaining the National Identity Card, leaving themselves **vulnerable** to persecution and unable to access services.*

13. *If a woman with this immigration status is the victim of domestic violence,(...) she is likely to be **fearful** of using either the criminal or civil justice system because she realises that she is at risk (...)*

Regarding the qualitative analysis of the data above, it is within this category where we find the clearest evidence that confirms our initial hypothesis. The figures obtained from the British corpus almost double those from the Spanish one as the former contains a much greater proportion of vocabulary items referring to the speaker's emotions or appealing to the listeners' or readers'. The contexts shown above instantiate that fact. As a consequence and also going back to the global results, it could be stated that a much greater proportion of terms expressing *affect*, in particular, and *appraisal*, on the whole, reinforces our interpretation of the data: the lesser presence of these vocabulary items in the Spanish corpus may reflect the constraints that this system imposes on legal actors, who, judging by the data presented above, seem to be less free to express stance towards textual content than their British counterparts.

5. CONCLUSION

This research has presented a corpus-based contrastive analysis of the expression of appraisal in two corpora of Spanish and British judicial decisions of roughly 3 million words each. To that end, the systemic linguistics model of *appraisal* was replicated in order to categorise the items of vocabulary identified as capable of expressing the attitude of the speaker/writer towards the content of a text.

In a first approach to the data, it was observed that the British corpus contained a higher proportion of such lexical items, namely, 1.08% against 0.72% (33% less) in the Spanish texts. When classified into categories, it was found that the proportions were similar with respect to the whole list of terms expressing appraisal (146 in Spanish and 274 in English). The categories (also known as systems) *appreciation* and *judgment* were distributed likewise in both corpora, where we found roughly 48% lexical items for the former and approximately 27% for the latter. The greatest differences were spotted within the category *affect*, which signals emotion on the part of the speaker/writer or appeals to the listener/reader. Roughly 20% of the words in the English vocabulary inventory belonged in this system, being twice as many as the Spanish set, which displayed 10%.

With regard to the research questions posed in the introduction, RQ1 reads as follows: to what extent can legal text reflect attitudinal stance? The answer would be clearly affirmative. It has been proved that, in spite of the common assumption that legal texts should leave little room for subjectivity, or at least, be as objective as possible (particularly judicial decisions), a considerable amount of lexical items expressing appraisal was found in two text collections of Spanish and British judicial decisions. As for the way in which this can be quantified, the

method employed to quantify these findings has been described throughout this study, where corpus linguistics techniques were applied for the contrastive analysis of legal discourse. To begin with, two type lists were obtained from two text collections of Spanish and British judgments, which allowed us to measure different parameters such as frequency or text distribution. Then, the two corpora were tagged and a closer linguistic profiling of the vocabulary inventory was carried out. After that, the lists of terms were categorised by replicating the linguistic model defined by appraisal theory, leading us to a closer and more systematic observation of such terms.

Concerning RQ2, which enquired about the way in which legal systems might condition the expression of interpersonal positioning in legal text, our interpretation of the findings, not surprisingly, has led us to conclude that judicial decisions emanating from civil law systems (like the Spanish one) display a greater tendency towards the inhibition of such positioning. On the other hand, those judgments issued in countries like the UK, where the legal system revolves around the principle of binding precedent, exhibit a tendency to express attitudinal stance more freely, probably due to the fact that the law is said to be judge-made, as opposed to civil law countries where it is codified.

To conclude, as an extension of the research presented in this paper, it would be very interesting to continue exploring the differences between civil and case law systems as regards the way in which their very nature conditions the speakers' attitude towards the propositional content of a text. Going beyond word limits and looking into more complex structures such as word sketches (Kilgarriff et al., 2014), that is, examining the implications of the use of certain grammatical patterns like the passive voice or the way in which verbs and objects colligate in both legal contexts might also cast light on those differences. This could be done by exposing the choice of specific lexical items which become the passive subject or the agent of verbs when dealing with controversial issues like irregular immigration. Furthermore, the contrastive analysis of the texts in both corpora by means of the examination of their syntax could complement the one presented above, not only providing a wider portrait of the differences in the expression of attitude across legal systems, but also offering a deeper characterisation of the phenomenon of immigration as seen through the lens of British and Spanish legal practitioners.

[Paper submitted 16 Sep 2019]

[Revised version received 21 Nov 2019]

[Revised version accepted for publication 27 Nov 2019]

References

- Alcaraz Varó, E. (1994). *El inglés jurídico: textos y documentos* [Legal English: Texts and documents]. Madrid: Ariel Derecho.
- Anthony, L. (2015). *TagAnt* (Version 1.1.0) [Computer software]. Tokyo, Japan: Waseda University.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse analysis and media attitudes: The representation of Islam in the British press*. Cambridge: Cambridge University Press.
- Biel, Ł., & Engberg, J. (2013). Research models and methods in legal translation. *Linguistica Antverpiensia*, 12, 1-11.
- Blinder S., & Allen W. L. (2016). Constructing immigrants: Portrayals of migrant groups in British national newspapers, 2010-2012. *International Migration Review*, 50(1), 3-40.
- Borja, A. (2000). *El texto jurídico inglés y su traducción al español* [Legal English and its translation to Spanish]. Barcelona: Ariel.
- Breeze, R. (2015). Teaching the vocabulary of legal documents: A corpus-driven approach. *ESP Today*, 3(1), 44-63.
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139-173.
- Eggs, S., & Slade, D. (1997). *Analysing casual conversation*. London: Casell.
- Finegan, E. (2010). Corpus linguistics approaches to legal language: Adverbial expression of attitude and emphasis in Supreme Court opinions. In M. Coulthard, & A. Johnson (Eds.), *The Routledge handbook of forensic linguistics* (pp. 65-77). London/New York: Routledge.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996-2005. *Journal of English Linguistics*, 36(1), 5-38.
- Goźdź-Roszkowski, S. (2017). Signalling sites of contention in judicial discourse: An exploratory corpus-based analysis of selected stance nouns in US Supreme Court opinions and Poland's constitutional tribunal judgments. *Comparative Legilinguistics*, 32, 91-115.
- Halliday, M. A. K. (2003). Introduction: On the architecture of human language. In J. J. Webster (Ed.), *On language and linguistics. Volume 3 in the collected works of M. A. K. Halliday* (pp. 1-32). London/New York: Continuum.
- Hunston, S. (2006). System and corpus: Two traditions with a common ground. In G. Thompson, & S. Hunston (Eds.), *System and corpus: Exploring connections* (pp. 1-14). London: Equinox.
- Kaltenbacher, M. (2006). Culture-related linguistic differences in tourist websites: The emotive and the factual. A corpus analysis within the framework of appraisal. In G. Thompson, & S. Hunston (Eds.), *System and corpus: Exploring connections* (pp. 269-292). London: Equinox.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The sketch engine: Ten years on. *Lexicography*, 1, 7-36.
- Marín Pérez, M. J., & Rea Rizzo, C. (2012). Structure and design of the BLRC: A legal corpus of judicial decisions from the UK. *Journal of English Studies*, 10, 131-145.
- Martin, J. R. (2003). Beyond exchange: APPRAISAL systems in English. In S. Hunston, & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142-175). Oxford: Oxford University Press.

- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. London/New York: Palgrave Macmillan.
- Mazzi, D. (2010). 'This argument fails for two reasons...': A linguistic analysis of judicial evaluation strategies in US Supreme Court judgments. *International Journal for the Semiotics of Law*, 23(4), 373-385.
- McEnery, T., & Baker, H. (2016). *Corpus linguistics and 17th-century prostitution: Computational linguistics and history*. London: Bloomsbury Academic.
- Mellinkoff, D. (1963). *The language of the law*. Boston: Little, Brown & Co.
- Orts, M. A. (2018). A bilingual, bicultural approach to detachment and appraisal in the law: Tracing impersonality and interaction in English and Spanish legal op-eds. *International Journal for the Semiotics of Law*, 31, 1-24.
- Pérez-Paredes, P., Aguado Jiménez, P., & Sánchez Hernández, P. (2017). Constructing immigrants in UK legislation and administration informative texts: A corpus-driven study (2007-2011). *Discourse and Society*, 28(1), 81-103.
- Pontrandolfo, G., & Goźdz-Roszkowski, S. (2014). Exploring the local grammar of evaluation: The case of adjectival patterns in American and Italian judicial discourse. *Research in Language*, 12(1), 71-91.
- Rothery, J., & Stenglin, M. (2000). Interpreting literature: The role of APPRAISAL. In L. Unsworth (Ed.), *Researching language in schools and communities: Functional linguistics perspectives* (pp. 231-264). London: Casell.
- Tiersma, P. (1999). *Legal language*. Chicago: The University of Chicago Press.
- White, P. R. R. (July 20, 1999). The language of attitude, arguability and interpersonal positioning. The appraisal website. Retrieved from <http://www.grammatics.com/appraisal/>

MARÍA JOSÉ MARÍN PÉREZ works as a lecturer at the English Department of the University of Murcia, Spain. She read her PhD on legal English and Corpus Linguistics in 2013, and has published articles in journals such as *Terminology*, *Corpora*, *International Journal of Language and Law*, *ASP*, *La Revue du GERAS*, amongst others. She has also attended conferences on ESP and Corpus Linguistics such as the Corpora and Discourse International Conference, at Lancaster University, several conferences organised by the Spanish Association of Languages for Specific Purposes (AELFE), and the Conference on Languages for Specific Purposes held at the University of Padova, Italy, in July 2019.