RiCL Research in Corpus Linguistics

Analysing hedging in legal discourse using small-scale and large-scale corpora

Holly Vass
Coventry University / United Kingdom

Abstract – Post-graduate law students are expected to handle a variety of legal and academic written genres. In these genres, meaning can be nuanced through a variety of means, including hedging. However, international law students often find decoding instances of hedging challenging, with interpretation of the writer's true meaning suffering as a result. The aim of this paper is to compare hedging in two written legal genres which are frequently used by post-graduate law students to underpin their own academic essays: Law review articles and Supreme Court decisions. To that end, two analyses were undertaken. The first was qualitative in nature and involved the identification of hedges in a small corpus of just under 50,000 words based on introspection and intuition as per Salager-Meyer (2000). The second analysis served to verify conclusions drawn by using WordSmith Tools 6.0 (Scott 2013) and a much larger corpus of nearly eight million words. Correlation was found between the results of the two analyses, indicating that qualitative results could be extrapolated to larger samples to some degree.

Keywords – hedging, legal genres, law review, Supreme Court, pragmatic competence, epistemic modality

1. Introduction

Postgraduate law students, those doing an LL.M. in the United States or United Kingdom for example, are required to handle a variety of legal and academic written genres during their studies. From a productive point of view, they could be asked to write case notes and academic essays. From a receptive point of view, they could be required to read and comprehend such diverse genres as journal articles, legal statutes, and law cases among others. This can be difficult for a native student who is the product of the target country's legal education and familiar with its legal system. For international students from other legal systems and with a different legal education background, this can be daunting.

This paper focuses on two legal written discourse genres which students would likely approach from a receptive perspective, but which could also be used to underpin their own academic essays, namely law review articles and Supreme Court decisions. Supreme Court decisions are public records of decisions made by the Supreme Court, the highest court in the U.S., and law review articles are scholarly research articles appearing in law journals, also known as law reviews. In particular, this paper aims to compare hedging in these two different legal genres as it is an area which is "notoriously problematic" for even advanced-level non-native speakers, but can be crucial in the interpretation of precedents which play a central role in the common law U.S. and U.K. legal systems (Abbhul 2006: 152). Hedging competence is part of pragmatic competence, or "the ability to communicate your intended message with all its nuances in any socio-cultural context and to interpret the message of your interlocutor as it was intended" (Fraser 2010: 15). However, the inability to hedge appropriately can result in second-language speakers being perceived as impolite, arrogant or offensive, and inability to interpret hedging can result in second language speakers failing to understand a native speaker's meaning (Fraser 2010: 15).

The WordSmith suite of concordancing tools (Scott 2013) was selected to explore the phenomenon of hedging in the corpus of law review articles and Supreme Court decisions as these tools allowed for lexico-grammatical items involved

in hedging to be readily identified. However, the use of concordancing tools also posed a dilemma. A text does not contain hedges *per se* but gets them through reader-writer interaction. Realisation of a hedge often responds to the writer's personal intention, while interpretation of a hedge requires introspection as well as understanding of context and shared background knowledge. The question was whether hedging could be validated by corpus research. Consequently, this research was conducted in two parts. Firstly, a qualitative textual analysis of a small corpus of seven texts was undertaken so that hedges could be identified in context, and a list of lexico-grammatical items signalling hedging was drawn up. Secondly, this list was used as the basis by which a much larger corpus could be explored using WordSmith Tools version 6.0.

2. HEDGING

The concept of 'hedging' has evolved alongside linguists' perception of language over the past 40 years. An early focus was on how lexico-grammatical forms can be used to nuance an utterance in order to better convey human perception of reality. Thus Lakoff (1973), in his landmark article 'Hedges: a study in meaning criteria and the logic of fuzzy concepts', explored how language can be used "to make things fuzzier or less fuzzy" (1973: 471), marking the first time 'hedging' was used as a linguistically oriented technical term. Lakoff cautioned early on, however, that "a far more sophisticated apparatus" than his would be needed to effectively study hedges (1973: 483).

As views on language evolved, discourse began to be seen as interactive in nature, and the relationship between both interlocutors, the addresser and the addressee, began to be analysed. Linguists became interested in how hedging not only affects semantic but also pragmatic meaning. The relationship between function and structure became more closely examined. As a consequence, the concept of hedging was further developed in pragmatics and discourse analysis to reflect its use within the context of social communication.

The concept of hedging was extended to other areas including epistemic modality, metadiscourse, stance, politeness, and vagueness. However, as the concept of hedging was extended to these other areas, its scope widened from Lakoff's original use and, in the opinion of many, it lost clarity in meaning (Markkanen and Schröder 1997). Nevertheless, in this paper hedging will be defined as a rhetorical strategy which functions on four levels: the interpersonal, the epistemic, the social, and the discourse community levels. On the interpersonal level, it shields the writer from potential disagreement and conflict by 'toning down' language so that a claim is presented as opinion rather than fact, thus affording the writer some protection from counter claims and helping to avoid potential conflict (Hyland 1998; Markkanen and Schröder 1997; Salager-Meyer 1997). On the epistemic level, it expresses uncertainty of knowledge and allows the writer to show commitment or lack thereof to the truth value of a proposition (Mauranen 1997; Salager-Meyer 1997). On the social level, it emphasises the subjectivity of a claim and allows the writer to recognise alternative viewpoints and open a "discursive space" in which these viewpoints can be put forth (Hyland 2005: 179). Finally, on the community level, it allows the writer to conform to a writing style which is accepted by the discourse community (Salager-Meyer 1997). Many authors see hedging as a unique phenomenon, which does not necessarily need to be linked to other phenomena and, as a result, several researchers have attempted to classify lexico-grammatical items used as hedges.

One problem with extending the concept of hedging is that, as Clemen (1997: 6) explains, "there is no limit to the linguistic expressions that can be considered as hedges ... no linguistic items are inherently hedges but can acquire this quality on the communicative context or co-text". Fraser (2010: 23) adds that hedging devices can be drawn from any syntactic category. Nevertheless, despite the problems with a comprehensive classification of what exactly constitutes a hedge, it is clear that there are certain central lexico-grammatical items which semantically embody the components of fuzziness, indeterminancy, and tentativeness (Varttala 2001: 34) and are generally interpreted as realisations of hedging by a majority of authors.

Hyland's (1996) taxonomy of hedging divides hedges into two categories: lexical and strategic. Thus, lexical hedges include modal verbs (*would*, *may*, *could*), epistemic lexical verbs (*indicate*, *suggest*, *appear*, *propose*), epistemic adjectives (*likely*, *possible*, *apparent*), epistemic adverbs (*apparently*, *probably*, *relatively*, *generally*), and epistemic nouns (*possibility*). Strategic hedges include other means of expressing tentativeness which do not fit in the category of lexical hedges, including the use of questions, conditional clauses and contrast markers, as well as a limited range of formulaic phrases. In Hyland's (1996) study on scientific research articles, lexical hedges represented approximately 85% of all items.

While studies of hedging in scientific and academic discourse abound, there appears to be far less research focusing on hedging in legal discourse specifically. Bhatia et al. (2004) explored the use of hedging in the legal problem question answer genre and concluded that both lexical surface hedges and non-lexical strategic hedges are "crucial for deductive reasoning and legal argumentation" (2004: 218). While Bhatia et al.'s (2004) focus was the use of hedging devices in the legal problem question answer, others have addressed hedging used in judicial opinions. Toska (2012) centred his study on the use of hedging to express epistemic attitude in Supreme Court judgments of the United Kingdom. He asserts that hedges convey a judge's stance on a given issue. Hinkle et al. (2012) looked at judgments at the U.S. District Court level

and found that "district court judges not ideologically aligned with the majority of the overseeing circuit judges use more hedging language in their legal reasoning in order to insulate these rulings from reversal" (2012: 407). Thus, the greater the ideological distance between the district court judge writing the judgment and the circuit court judge reviewing his decision, the greater the amount of hedging used.

3. Methodology

This study sets out to examine the frequency and role of lexical hedging in two written legal genres, American law review articles and U.S. Supreme Court decisions. The corpora used are specialised and limited not only in terms of genre but also time (written between 1998 and 2003) and place (the United States), reflecting a desire to better situate the data collected in a particular context. Two different analyses were carried out. The first analysis was purely qualitative and, to facilitate data gathering, a small but representative corpus of seven total texts was used: three law review articles and four U.S. Supreme Court decisions, totalling 30,000 and 18,000 tokens, respectively. The second analysis was quantitative and used much larger corpora of the two genres, totalling around 4 million tokens each.

Salager-Meyer (2000) advocates an approach that makes use of introspection on the part of the researcher when investigating hedging, due to the role hedging competence plays in the identification of a hedge. Thus, in the first analysis, a qualitative method, namely textual analysis, rather than quantitative data gathering and interpretation techniques, was used. This was made feasible by using such a small, but representative corpus as gathering qualitative data using a large-scale corpus could be "tedious, inaccurate and incomplete" (Bhatia et al. 2004: 212).

Accordingly, the seven representative texts were read, and hedges were identified. After their identification, lexico-grammatical items signalling hedges were coded according to Hyland's (1996) taxonomy. Thus, items were separated into the following categories: modal verbs, epistemic lexical verbs such as *seem* and *appear*, epistemic adverbials such as *somewhat* and *potentially*, epistemic adjectives such as *likely*, *possible* and *debatable*, and epistemic nouns such as *assumption*. Finally, a preliminary list of potentially high frequency items used to signal hedging was drawn up, and differences and similarities between the genres were noted. Initial conclusions were reached concerning realisation of hedging in the two genres.

The challenge of the present study was to support initial conclusions based on a very small and limited corpus with empirical data gathered from a larger corpus through the use of corpus-based analyses tools, such as concordancing software. This need to balance a qualitative and a quantitative approach led to a second analysis.

The second analysis made use of a more comprehensive list of 130 potentially high-frequency lexical hedges drawn from findings in the first analysis as well as examination of the literature (Bhatia et al. 2004; Hyland 1996; Salager-Meyer 1997) (see Table 1). Bhatia et al.'s (2004) original list was refined somewhat as their list included lexical items that can indicate not only uncertainty, but also certainty (*must*, *cannot*, *conclude*, *confirm*, *conclusion*, *requirement*, *obvious*, *definite*, *certainly*, *definitely*). These words expressing certainty were eliminated as they do not express a hedge.

MODAL VERBS						
could	might	may	ought	should	would	
EPISTEMIC LEXICAL VERBS						
allege	appear	argue	assume	believe	claim	consider
depend	doubt	estimate	expect	indicate	infer	(not) know
(it is (not)) known	note	predict	presume(d)	propose	report	seem
speculate	suggest	tempt	tend	think		
EPISTEMIC NOUNS						
allegation	argument	assumption	belief	(un)certainty	chance	estimate
impression	likelihood	(im)possibility	(im)probability	mis-emphasis	suggestion	
EPISTEMIC ADJECTIVES						
alleged	apparent	arguable	(un) (not) certain	(un) (not) clear	debatable	doubtful
general	imaginable	(un) likely	(im)possible	potential	(im)probable	purported
questionable	(un)reasonable	so-called	unclear	virtual		
EPISTEMIC ADVERBS						
allegedly	almost	apparently	approximately	arguably	essentially	generally
ideally	in that	maybe	partly	per se	perhaps	possibly
potentially	presumably	presumptively	probably	relatively	roughly	somewhat
somehow	sometimes	theoretically	typically	usually	whenever	virtually

Table 1. Enlarged list of potential lexical hedges

In addition, words which could express a hedge in general English, but more often in legal English have a specific, technical meaning (*evidence*, *alleged*, *question*, *claim*, *reasonable*, *doubt*) were also eliminated. Only epistemic modal verbs expressing some sort of doubt or lack of commitment were included.

Not only was an enlarged match list of potential hedges used in the second analysis, but also a much larger corpus of articles and opinions, consisting of roughly four million tokens for each genre thus reaching a total of approximately eight million tokens. The enlarged list of potential hedges was lemmatised and loaded onto the WordList function of WordSmith Tools 6.0 (Scott 2013) as the match list file. Results were then colour-coded according to part of speech. In addition, the Concord function of WordSmith Tools 6.0 was used for checking the context of random potential hedges to verify if they were being used as such, and for extracting examples.

The second analysis allowed for a more complete picture of potential hedges to emerge. Baker (2010: 94) lists three clear advantages of corpus-based analyses. Firstly, concordancing software programs such as WordSmith Tools give researchers access to linguistic patterns and trends quickly and accurately. This allows researchers to confirm or refute hypotheses that would have otherwise been impossible. Finally, since concordancing tools allow researchers to quantify linguistic patterns, more solid conclusions about language can be reached.

4. RESULTS

Initial reading of the corpora in search of lexico-grammatical items which could potentially signal a hedge suggested that hedging was realised in the two genres differently in terms of both range of items as well as their frequency of use. This initial exploration of the small corpora allowed for certain trends and tendencies to be postulated. In general, hedging was far more common in law review articles than in Supreme Court opinions. Both incidence of hedging and lexico-grammatical items used to signal hedging in the corpus of U.S. Supreme Court opinions examined appeared to be much more limited when compared to the corpus of American law review articles.

Moreover, while both modal and lexical verbal hedging could be found in the corpus of law review articles studied, there were few instances of verbal hedging of any type in U.S. Supreme Court opinions. In fact, though three modal verbs are present in the corpus of U.S. Supreme Court opinions, actual incidence of epistemic modality was extremely restricted.

Furthermore, in addition to lexical and modal verbs, other lexical items, most notably adverbs and adjectives, also signalled hedging in the law review articles. By contrast, instances of non-verbal hedging in the U.S. Supreme Court opinions were generally signalled in one of three ways: prepositional phrases, adverbial clauses introduced by the concessive coordinators *while* and *although*, or negation. Prepositional phrases were commonly formed by both simple (*for*, *in*, *to*) or complex (especially two-word eg. *due to*) prepositions whose complement was a noun which could be modified by an attenuating adjective such as *certain* or *limited*:

> In Montana v. United States, 450 U.S. 544 (1981), we held that, *with limited exceptions*, Indian tribes lack civil authority… Atkinson Trading Co. v. Joe Shirley, et al., 532 U.S. 645 (2001)

Four of the prepositional phrases (*for purposes of the appeal*, *in the context of safety and administrative regulations*, *in certain limited circumstances*, and *with limited exceptions*) used in Supreme Court opinions seemed to serve the function of limiting the context or the circumstances in which a rule or decision might apply. Similar prepositional phrases did not appear in law review articles.

Sentence initial clauses beginning with a concessive subordinator were used in conjunction with negation to present two contrasting ideas. In the following example, the justices claim not to dispute the basis of some previous approaches to a question of law; however, they seem to blame 'the statute' for not letting common sense prevail:

> *While we do not dispute the common sense of this approach*, the words of the statute do not permit it. Caron v. United States, 524 U.S. 308 (1998)

In summary, examination of the results of the first analysis allows for some tentative conclusions to be reached:

1. There appears to be a higher incidence, in general, of hedging in law review articles than in Supreme Court opinions.
2. Lexical hedging seems more common and more varied in law review articles, being achieved through the use of a range of modal verbs, epistemic lexical verbs, epistemic adjectives, and epistemic adverbs.
3. Lexical hedging seems much more limited in Supreme Court opinions, both in terms of range and frequency.
4. Strategic hedging appears to occur more commonly than lexical hedging in Supreme Court opinions, and is achieved by means of clauses initiated by concessive subordinators, prepositional phrases, and negation.
5. Strategic hedging appears to be more limited in range in law review articles than in Supreme Court opinions.

While certain conclusions could be proposed based on this initial qualitative analysis, one issue was that such an analysis is very limited in terms of the size of corpora used. Another issue to be addressed was that a qualitative analysis of this type relies purely on the researcher's introspection for data gathering which, while important in the identification

of hedges (Salager-Meyer 1997), introduced an element of subjectivity into the analysis. The validity of these initial, tentative conclusions required verification using a larger, more representative corpus, as well as quantitative methods which could limit the degree of subjectivity inherent in any qualitative research. The question would be if a quantitative analysis of larger corpora using concordancing software would bear the same results. The following table shows the most frequent lexical verbs found in the law review articles and Supreme Court decisions, respectively.

Modal verbs Lexical verbs Adjectives Adverbs Nouns

Law review articles

N	Word	Freq.
1	WOULD	14180
2	MAY	10657
3	SHOULD	5662
4	COULD	5515
5	MIGHT	5371
6	ARGUE	3245
7	SUGGEST	3088
8	LIKELY	2778
9	GENERAL	2756
10	CONSIDER	2654
11	SEEM	2279
12	THINK	2185
13	ASSUME	1721
14	POSSIBLE	1647
15	BELIEVE	1609
16	EXPECT	1595
17	POTENTIAL	1505
18	PERHAPS	1344
19	APPEAR	1317
20	PROPOSE	1316
21	SOMETIMES	1133
22	DEPEND	1008
23	POSSIBILITY	900
24	RELATIVELY	820
25	PROBABILITY	807
26	TEND	744
27	INDICATE	719
28	ALMOST	700
29	ASSUMPTION	650
30	ATTEMPT	559
31	TYPICALLY	554
32	USUALLY	485

Supreme Court decisions

N	Word	Freq.
1	WOULD	9963
2	MAY	6897
3	GENERAL	4369
4	COULD	3810
5	SHOULD	3046
6	ARGUE	2898
7	CONSIDER	2341
8	MIGHT	2076
9	SUGGEST	1442
10	ALLEGE	1340
11	BELIEVE	1340
12	THINK	1203
13	ASSUME	963
14	APPEAR	887
15	LIKELY	885
16	INDICATE	691
17	SEEM	676
18	POTENTIAL	624
19	POSSIBLE	615
20	UNREASONABLE	614
21	PROPOSE	588
22	POSSIBILITY	510
23	DEPEND	509

Table 2. Lexical hedges in the corpus

Upon examination of the tables, it can be seen that modal verbs continued to play a significant role in hedging in both genres. In addition, epistemic adjectives appeared to be involved in hedging as well, especially *likely*, *general*, *possible*, and *potential*. However, while in the previous analysis it appeared that lexical verbs did not play a role in hedging in Supreme Court decisions, here it is apparent that they were important in both genres. Epistemic adverbials continued to play a more significant role in law review articles than in Supreme Court decisions, with the notable exception of *perhaps*.

A comparison of number of hedges of each type yielded the results in Table 3. Hedging appeared to be more widespread in both genres than was apparent in the first quantitative analysis. If it was originally thought that there were 84 hedges per 10,000 words in law review articles, and 56 hedges per 10,000 words in Supreme Court decisions, the table below shows 172 and 128 hedges per 10,000 words respectively.

Hedge type	Law review articles		Supreme Court decisions	
	Per 10,000 words	Raw total	Per 10,000 words	Raw total
Modal verbs	87	41,385	68	25,792
Lexical verbs	50	24,039	39	14,878
Adjectives	18	8,686	19	7,107
Adverbs	12	5,516	1	385
Nouns	5	2,357	1	510
Total	172	81,983	128	48,672

Table 3. Hedges per 10,000 words in the corpus

This is undoubtedly due to a greater number of potential hedges being sought and may lead to the conclusion that hedging could be considered even more widespread if more lexical items were taken into account. Still, there appears to be approximately 25% more hedging in law review articles than in Supreme Court decisions.

The second analysis appeared to corroborate the three tentative conclusions reached after a first qualitative analysis of a smaller corpus. First, overall, there was a 25% higher incidence of lexical hedging in law review articles than in Supreme Court decisions. Second, not only was lexical hedging more common in law review articles, but also it was more varied, being achieved through a range of modal verbs, epistemic lexical verbs, epistemic adjectives, and epistemic adverbs. The one exception is epistemic adjectives, which appears to be slightly more frequent in Supreme Court opinions. Lexical hedging seems much more limited in Supreme Court opinions, both in terms of range and frequency. This is represented visually in Figure 1.

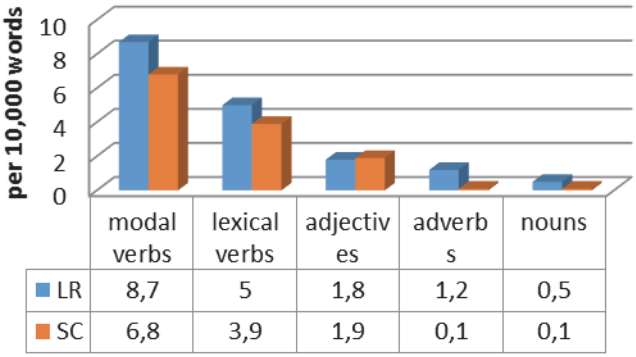

Figure 1. Frequency of types of hedges in the corpora, where LR = law review articles and SC = Supreme Court decisions

Third, both genres appeared to rely heavily on modal verbs to convey uncertainty or lack of commitment as to the truth value of an utterance. However, proportionally, as illustrated in the pie charts below (Figure 2), modal verbs, lexical verbs, and adjectives appeared to play a more significant role in Supreme Court decisions than in law review articles, while the opposite was true when considering adverbs and nouns.

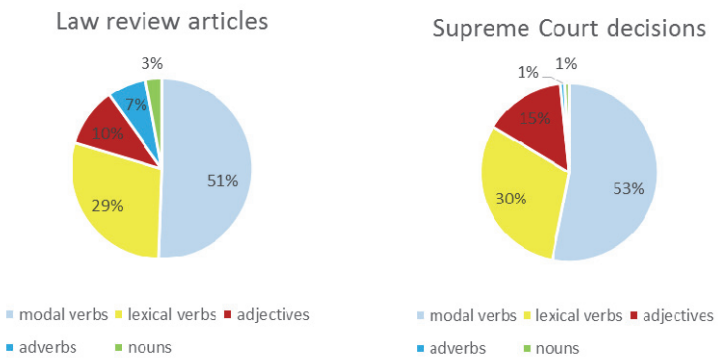

Figure 2. Comparison of lexical hedges in terms of percentage of use

Omitting such nouns as *evidence*, *claim*, *question*, and *doubt* clearly influenced the results of the second analysis, as only 3% of all hedges in the law review articles and 1% of all hedges in the Supreme Court opinions were nouns.

In addition to lexical hedges, two strategic hedges were also analysed. These were *if*-clauses, indicating a condition, and the contrast markers *while*, *whereas*, and *although*. Figure 3 compares the incidence of these hedges in the genres per 10,000 words.

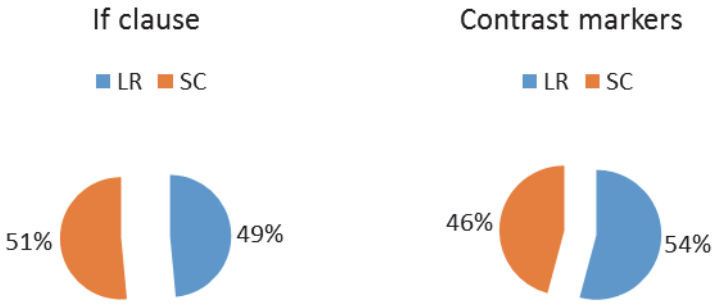

Figure 3. Strategic hedges

5. Discussion

This study has examined the phenomenon of hedging in both a small corpus of 48,000 tokens as well as a large corpus of over eight million tokens. The rationale was to verify conclusions drawn from a qualitative analysis of the small corpus by means of a quantitative analysis of a much larger corpus using WordSmith Tools 6.0 with a view to determining if there were substantial differences between the two analyses.

Initial results suggested by the qualitative analysis thus became predicted results of the larger, quantitative analysis. The first predicted result was that there would be more lexical hedging in law review articles than in Supreme Court decisions, especially through modal verbs. The quantitative analysis verified this and indicated that there was nearly 25% more lexical hedges in law review articles, and modal verbs were shown to be the main source of tentativeness in both genres. These results can be explained by looking at the communicative purposes attributed to each genre. Hedging aids the writer in presenting a claim as opinion rather than fact which has the dual effect of toning down the writer's claim so that it is not overly assertive, which is generally frowned upon by the discourse community, and opening a discursive space in which other viewpoints are recognised. This is more important in law review articles which, in addition to informing the reading public on issues of import, also have the purpose of providing a forum for discussion between members of the discourse community. Article writers also stand a better chance of gaining acceptance of their views within a body of work if deference and modesty are conveyed, and if their views are accepted, their status within the field will be elevated.

A second predicted result was that there was a wider variety of lexical hedges in law review articles than in Supreme Court decisions. This was also verified by the quantitative analysis of the larger corpus. Indeed epistemic nouns and adverbs were very low in frequency in Supreme Court decisions. However, as nouns and adverbs played an insignificant role in Supreme Court decisions, modal verbs, lexical verbs, and adjectives gained in importance in the genre. In fact, the new analysis pointed to a greater frequency of epistemic adjectives in Supreme Court decisions than in law review articles.

A third predicted result was that there would be a higher incidence of strategic hedges in Supreme Court decisions than in law review articles. The quantitative analysis indicated that this was true in the case of *if*-clauses, but not in the case of contrast markers. *If*-clauses could be used more frequently in Supreme Court decisions than in law review articles where they serve to strengthen the justices' argument by weakening a claim to the contrary, presenting it as unlikely or impossible. In addition, counterfactual conditionals are strongly related with causation, the relationship between a cause and effect, and defining this relationship is important in settling a legal dispute. As one of the main communicative purposes of a Supreme Court decision is to settle the dispute at hand, entertaining a variety of hypothetical situations aids the justice in providing a rationale for the holding. As another communicative purpose is to clarify a point of law for future cases, hypotheticals aid future trial lawyers who may find themselves involved in a case with similar circumstances.

The second quantitative analysis using WordSmith Tools 6.0 was used to verify the results of the qualitative analysis on the small corpus. However, while giving a more detailed picture of hedging in the two genres, use of a large corpus of nearly eight million tokens entailed such extensive results that it became impossible to verify if a potential hedge was being used as exactly that in context. One criticism of corpus-based analyses using concordancing software is that they do not always take into account the communicative context in which a sample of language is created (Widdowson 1998). This is particularly problematic when analysing pragmatic features of a text, which necessarily must be related to

the socio-cultural context (Flowerdew 2005). Modal verbs, for instance, present a challenge. For example, *may* and *might*, according to Hyland (1998: 116), are often "considered prototypical hedges", and indeed *may* is the second most frequent modal verb used in both genres with a total of just over 17,500 instances. But, perhaps *may* occurs so frequently in the corpora because it is a modal verb which appears to be used more frequently in legal English than in general English (Foley 2001: 193). This is because in legal English *may* can have a very specific deontic meaning of "conferring discretionary power, either directly or indirectly" (Williams 2007: 121):

> For representation of a benefits claimant at the administrative level, an attorney *may* file a fee petition or a fee agreement. (535 U.S. 789 (2002))

Nevertheless, the use of the Concord function of WordSmith tools showed that in both of the genres under study epistemic *may* was also apparent, as in the following example from a Supreme Court decision, in which *may* signals a possibility:

> No matter how odd or deficient trial counsel's performance *may* seem, that lawyer *may* have had a reason for acting as he did.... Or it *may* turn out that counsel's overall performance was sufficient despite a glaring omission. (538 U. S. 500 (2003))

Yet, when there are so many instances of these items, it is not always feasible to check each individual context to determine how it is being used. Validity of results can suffer as a consequence.

6. Concluding remarks

A better understanding of the most frequent lexico-grammatical items used as hedges could potentially focus academic English classes and benefit a wide range of non-native students in higher education. Studies have shown that different cultures approach hedging in different ways. For example, some academic cultures deem categorical statements as more persuasive (Tessuto 2011) and interpret hedging as "negative, evasive concepts" (Alonso et al. 2012: 47) which obscure the writer's claims and indicate insecurity. A more complete analysis of hedging could go some way to change students' misconceptions.

However, while both analyses undertaken in this study have proved insightful, in order to assure validity of results it would be necessary to make one further and important change to the research methodology. The list of potential hedges to serve as the match list in the WordList function of WordSmith tools is now set, but the current corpus, at over eight million tokens, is simply unmanageable. The corpus could be reduced to just over one million tokens which would allow it to retain its representativeness while at the same time allow for verification of the true use and nature of each potential hedge in context.

References

Abbuhl, Rebekha. 2006. Hedging and boosting in advanced-level L2 legal writing: the effect of instruction and feedback. In Heidi Byrnes, Heather D. Weger-Guntharp and Katherine A. Sprang eds. *Educating for advanced foreign language capacities: constructs, curriculum, instruction, assessment*. Washington, DC: Georgetown University Press, 152–164.

Alonso Alonso, Rosa, María Alonso Alonso and Laura Torrado Mariñas. 2012. Hedging: an exploratory study of pragmatic transfer in nonnative English readers' rhetorical preferences. *Ibérica* 23: 47–64.

Baker, Paul. 2010. Corpus methods in linguistics. In Lia Litosselti ed. *Research methods in linguistics*. London: Continuum, 93–116.

Bhatia, Vijay K., Nicola M. Langton and Jane Lung. 2004. Legal discourse: opportunities and threats for corpus linguistics. In Ulla Connor and Thomas Upton eds. *Discourse in the professions: perspectives from corpus linguistics*. Amsterdam: John Benjamins, 203–231.

Clemen, Gudrun. 1997. The concept of hedging: origins, approaches and definitions. In Raija Markkanen and Hartmut Schröder eds. *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts*. Berlin: De Gruyter, 235–248.

Flowerdew, Lynne. 2005. An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes* 24: 321–332.

Foley, Richard. 2001. Going out in style? *Shall* in EU legal English. *Proceedings of the Corpus Linguistics Conference 2001*. Lancaster: University of Lancaster, 185–195.

Fraser, Bruce. 2010. Pragmatic competence: the case of hedging. In Gunther Kaltenböck, Wiltrud Mihatsch and Stefan Schneider eds. *New approaches to hedging*. New York: Academic Press, 15–34.

Hinkle, Rachael K., Andrew D. Martin, Jonathan David Shauband and Emerson H. Tiller. 2012. A positive theory and empirical analysis of strategic word choice in district court opinions. *Journal of Legal Analysis* 4/2: 407–444.

Hyland, Ken. 1996. Talking to the academy: forms of hedging in science research articles. *Written Communication* 13/2: 251–281.
Hyland, Ken. 1998. *Hedging in scientific research articles*. Amsterdam: John Benjamins.
Hyland, Ken. 2005. Stance and engagement: a model of interaction in academic discourse. *Discourse Studies* 7/2: 173–192.
Lakoff, George. 1973. Hedges: a study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 2: 458–508.
Markkanen, Raija and Harmut Schröder. 1997. Hedging: a challenge for pragmatics and discourse analysis. In Raija Markkanen and Hartmut Schröder eds. *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts*. Berlin: Walter de Gruyter, 3–18.
Mauranen, Anna. 1997. Hedging in language revisers' hands. In Raija Markkanen and Hartmut Schröder eds. *Hedging and discourse: approaches to the analysis of a pragmatic phenomenon in academic texts*. Berlin: Walter de Gruyter, 115–133.
Salager-Meyer, Francoise. 2000. Procrustes' recipe: hedging and positivism. *English for Specific Purposes* 19: 175–187.
Salager-Meyer, Francoise. 1997. I think that perhaps you should: a study of hedges in written scientific discourse. In Tom Miller ed. *Functional approaches to written text: classroom applications*. Washington, DC: English Language Programs-United States Information Agency, 105–118.
Scott, Mike. 2013. *WordSmith Tools 6.0* <http://www.lexically.net/wordsmith>.
Tessuto, Girolamo. 2011. Legal problem question answer genre across jurisdictions and cultures. *English for Specific Purposes* 30: 298–309.
Toska, Bledar. 2012. Epistemic hedges and boosters as stance markers in legal argumentative discourse. *Topics in linguistics: approaches to text and discourse analysis* 10: 57–62.
Varttala, Teppo. 2001. *Hedging in scientifically oriented discourse. Exploring variation according to discipline and intended audience*. *Acta Electronica Universitatis Tamperensis* 138. Tampere: University of Tampere.
Widdowson, Henry G. 1998. Context, community and authentic language. *TESOL Quarterly* 32/4: 705–716.
Williams, Christopher. 2007. *Tradition and change in legal English: verbal constructions in prescriptive texts*. Bern: Peter Lang.