



RENÁTA PANOCOVÁ AND LUKÁŠ LUKAČÍN

## EPISTEMIC MODAL MARKERS IN TWO DOMAINS OF ACADEMIC RESEARCH PAPERS IN ENGLISH

### Abstract

In this paper we compare the combinations of epistemic modal markers in research articles in medicine and humanities written in English. In our analysis, we focus on three aspects. First, we look at the distribution of combinations of MODAL AUXILIARY + MODAL ADVERB (emphasizer) used in an epistemic sense in two subdomains of academic writing, ACAD: Medicine and ACAD: Humanities in COCA (Davies, 2008–2018). Second, we investigate the statistical significance of the differences between the two subdomains. Finally, we consider the relevance of the epistemic modal markers in presenting the argumentation line in research articles in medicine and humanities. The results demonstrate the difference in preference in co-occurrences of the selected modal markers in the two distinct academic subcorpora and indicate to what extent they are a significant feature to be included in developing academic writing skills, which is crucial for the effective and convincing communication of research results.

### Keywords

*Epistemic modality; modal auxiliaries; lexical modal markers; corpus analysis*

---

### 1. Introduction

In recent years, there has been an increasing interest in studying epistemic modality in academic texts. The traditional view of *epistemic modality* is formulated by Nuyts (2016: 38) as involving “an estimation, [...] typically but not necessarily by the speaker, of the chances or the likelihood that the state of affairs expressed in the clause applies in the world”. This means that epistemic modality helps us

understand what information is certain or uncertain, whether it is a fact or belief. Grammatical expressions for epistemic modality include epistemic modal auxiliaries, clitics, or affixes (e.g. Palmer 1986, 2001; Bybee et al. 1994; Boye 2012, 2016). In addition, many languages have lexical epistemic modal expressions such as verbs, adjectives, adverbs, and nouns (Boye 2016: 118–122). The situation in English is illustrated in (1).

- (1) a. This **may** be a conceivable explanation. [ACAD: Humanities, COCA]  
 b. Programs for music teachers **will clearly require** considerably more depth of musical expertise than those for early childhood teachers.  
 [ACAD: Humanities, COCA]

In (1a) there is an example of a modal auxiliary. Modal auxiliaries in English represent a closed set. The epistemic modal meaning can be paraphrased as that the speaker admits that perhaps it is possible to accept the explanation. In (1b) we can see that epistemic modal meaning is expressed not only grammatically, by the modal auxiliary *will*, but also by the adverb and the lexical verb. Aarts (2011: 307) includes *require* to the class of modal lexical verbs which take a direct object, as in (1b). In addition, nouns and adjectives can also have a modal meaning. The effect is intensified by the use of combinations of several modal markers as in (1b).

One of the domains where epistemic modality plays an important role is scholarly writing. It is generally known that at present science is marked by a rapid progress. This is reflected in the necessity of a convincing presentation of research findings in scientific journal articles. A number of studies of the rhetorical style of research articles (e.g. Myers 1990; Hyland 1998) demonstrate that writing a scientific text “involves selective representation and rhetorical reconstruction as a means of anticipating negative responses to claims” (Hyland 1998: 18). In order to achieve this, the use of epistemic modal markers is crucial. In line with Pérez-Llantada (2010: 25–26) “epistemic modality has proved to be a highly routinised phenomenon in academic writing, yet rhetorically variable across cultural contexts”.

Many studies investigated the use of epistemic modality as a rhetorical feature for reporting claims in academic argument in research papers written in English and by native speakers of English (e.g. Salager-Meyer 1994; Skelton 1997; Hyland 1996, 1998, 2001, 2012; Thompson 2001; Varttala 2003; Vold 2006; Ardizzone and Pennisi 2012; Chovanec 2012). The main reason is that English has taken the position of *lingua franca* in communicating research results. Large (1983: 18) reports that already in 1980, 85 per cent of research articles in biology and physics were written in English, in medicine slightly less with 73 per cent. Maher (1986) found almost identical results in the field of medicine. This tendency of dominant use of English in medical research papers continues. Giannoni (2008) reports, for instance, that more than 99 per cent of medical research papers by Italian authors are in English and Gunnarson (2009) describes similar findings in

Scandinavian countries. According to Crystal (2012: 112), a significant increase of research articles written in English is observed “even in a language-sensitive subjects such as linguistics, where in 1995, nearly 90 per cent of the 1,500 papers listed in the journal *Linguistic Abstracts* were in English”.

These observations triggered further investigation of similarities and differences in the use of epistemic modality in research articles written by native speakers of English and non-native authors. The results demonstrate that there has been a tendency to adopt a normative scientific writing style used in English among non-native authors for the sake of international recognition and sharing the knowledge (e.g. Fergusson 2007; Flowerdew 2007). Ngula (2015, 2017) reports that Ghanaian scholars considerably underuse epistemic modal verbs in their research papers written in English. Vold (2006) compares the use of epistemic modal markers in linguistics and medicine in three languages, French, English, and Norwegian. She found out that the differences were significant and concluded that non-native English authors tended not to follow the conventional use of epistemic modality typical of scientific style in English. Orta (2010) and Pérez-Llantada (2010) arrive to a similar conclusion for Spanish scholars, and Panocová (2008) for Slovak researchers.

Another direction of the research into epistemic modality concerns its use in different registers. Kranich and Gast (2015) investigated the distribution of epistemic modal markers in four main register types (science and popular science, press, general fiction and biographies, essays, belle lettres, and popular fiction) in British English and American English. They focused on frequencies of modal auxiliaries (*can, may, must, might, could*), lexical modal verbs (e.g. *seem, appear*), modal adjectives or adverbs (e.g. *probably, perhaps*), and modal periphrasis (e.g. *I doubt if*). The results of their analysis of mixed-genre corpora from 1961 and those from 1991 indicate that “there is a significant increase in the use of modal auxiliaries for the expression of epistemic modality at the expense of lexical markers” (Kranich and Gast 2015: 4). These observations differ from earlier studies by Holmes (1983) and Hermerén (1986) who both claim that in their data modal items from other classes, especially lexical verbs and adverbs, are much more frequent with modal meanings than the closed class modal auxiliaries.

Against this background, this paper aims to investigate the way co-occurring modal markers can operate in union to reinforce the expressed modal meaning (Huddleston and Pullum, 2002), in particular between emphasizees such as *actually, clearly, of course*, etc. (Quirk et al. 1985; Hoye 1997) and the generally established set of modal auxiliaries. The co-occurrences of the two categories will be investigated in two academic subcorpora of The Corpus of Contemporary American English (COCA) (Davies, 2008-2018), namely ACAD: Medicine and ACAD: Humanities. In this paper we address the three main research questions:

1. What is the distribution of epistemic combinations MODAL AUXILIARY + MODAL ADVERB (emphasizer) in the two subdomains of academic domains?
2. How significant are individual combinations MODAL AUXILIARY + MODAL ADVERB (emphasizer) in ACAD: Humanities and ACAD: Medicine?
3. How relevant are these epistemic modal markers in presenting the argumentation line in scientific articles in medicine and humanities?

The paper is organized in five sections. After the background motivation for the present study in this section, the theoretical treatment of epistemic modality is discussed in Section 2. Methodological considerations of data collection are described in Section 3. The results of the analyses are given in Section 4. The main findings are summarized in the Conclusion.

## 2. Theorizing about epistemic modality

As already mentioned in Section 1, epistemic modality is linked to the speaker's evaluation of possibilities and it frequently "indicates the speaker's confidence (or lack of confidence) in the truth of the proposition expressed" (Coates 1983: 18). This is illustrated by the examples in (2).

- (2) a. Care homes **will certainly** have to diversify and specialise to handle the impact of public funding cuts. [ACAD: Medicine, COCA]  
 b. That is, the quality of students' practice **may actually** be more important than the amount of time spent practicing. [ACAD: Humanities, COCA]

It is immediately obvious that the example sentences in (2) come from scientific journal papers. In (2a) the authors use *will* to evaluate possibilities, most likely on the basis of an analysis of the present situation in care homes management and organization. In addition, the scholars are confident in their prediction of the direction of future changes. Their confidence is stressed by the use of the modal adverb *certainly* in the role of an emphasizer. In (2b) the possibilities are assessed with less force than in (2a). The modal meaning is that perhaps it is possible that quality of the practice is more essential than its quantity. Similarly, the modal adverb in (2b) has an emphasizing function.

Nuyts (2016: 44–48) discusses characteristic features and properties of modal categories. A relevant dimension of epistemic modality is the distinction between "subjective" and "objective" modal category (cf. Lyons 1977; Coates 1983; Palmer 1986; Verstraete 2001; Nuyts 2016). Subjective epistemic modality relies on a purely subjective assessment of a situation as true or false. In contrast, objective epistemic modality "expresses an objectively measurable chance that the

state of affairs under consideration is true or not” (Nuyts 2016: 45). This can be illustrated in (3).

- (3) Jane **must** be at school by now.

The example in (3) may be interpreted on the basis of the speaker’s logical inference. It is obvious that the speaker assumes that Jane is at school at the moment, for instance, because s/he knows when Jane left and how long the journey to school takes. What is important is the speaker’s subjective evaluation, often without explicit statements of the facts underlying it. The alternative interpretation of (3) may be that the speaker wants to indicate that there is a computable chance that Jane is at school. In the light of what is known, i.e. objective facts, we can be sure that Jane is at school. This is labelled as objective epistemic modality. In the Czech linguistic tradition, a similar distinction between *objective modality* and *subjective modality* is maintained, as discussed in detail by Ševčíková (2009: 5–16).

According to Coates (1983: 18), objective epistemic modality is a primary concern of logicians where it is labelled as *alethic modality*, which was introduced by von Wright (1951). Nuyts (2016: 38) sees a similarity between the distinction between subjective and objective epistemic modality in linguistic semantics and the distinction between alethic and epistemic modality in modal logic. Interestingly, the term *alethic modality* is almost never used in linguistics. Some linguists do not accept this notion at all. Palmer (1986: 11) makes it explicit that no distinction can be drawn between what is logically true and the speaker’s belief of what is true.

Epistemic modality is often considered to be of a scalar nature. For instance, Nuyts (2001, 2016) states that the treatment of epistemic modality as a matter of degree is generally accepted in functional linguistics. Given perhaps the most common subjective interpretation of epistemic modality, the relevant modal markers can be placed between two ends of a continuum, confidence and doubt. Hermerén (1978) gives a single-scale model, which contrasts with two parallel scales in Coates’s model (1983: 19) in Figure 1.

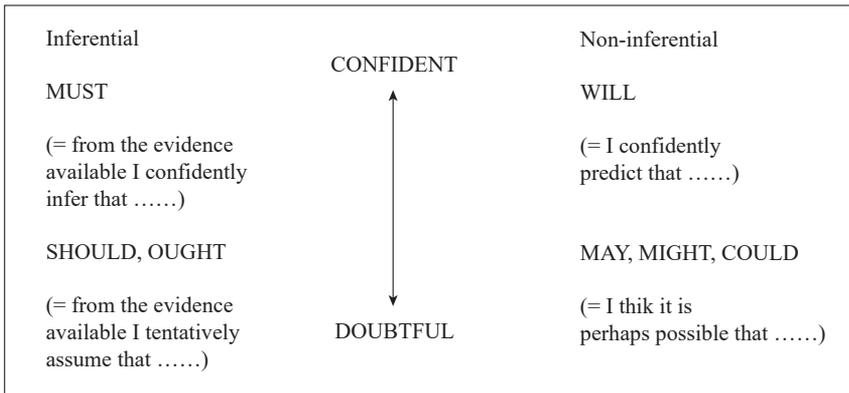


Figure 1. A two-scale model of epistemic modals, adapted from Coates (1983: 19)

The model in Figure 1 includes inferential and non-inferential elements. This feature makes it possible to differentiate between the confidence of the speaker based on the information s/he has and the one based on intuition and conviction not necessarily based on external evidence. This model takes into account the dimension of subjectivity and objectivity as described above.

Modal auxiliaries frequently co-occur with adverbs as presented in (1b) and (2). It was clearly demonstrated that these adverbs have a reinforcing effect on the truth-value of the clause which they apply to. In (1b) and (2a) the adverbs focalize the modal auxiliary and determine the degree of confidence of the speaker in relation to his/her prediction. Adverbs of this type belong to the subjunct class of adverbials, more precisely to emphasers (e.g. Bolinger 1972; Hoyer 1997; Quirk et al. 1985; Huddleston and Pullum 2002).

Hoyer (1997: 157–158) distinguishes two classes of emphasers. Type A includes emphasers concerned with the speaker's affirmation of truth, e.g. *actually, certainly, clearly, definitely*. Type B covers emphasers concerned with the speaker's value judgement, e.g. *frankly, honestly, literally, fairly*. According to Hoyer (1997: 157), the emphasers differ in the way they interact with modals, e.g. *certainly* has an inherently modal meaning and therefore emphasizes the truth-value of the sentence. Emphasers are not a closed class. For instance, although *probably* in (1b) is not listed among Type A emphasers by Hoyer, it certainly falls in this category. This also means that their list is not exhaustive and there are often differences among different authors. Emphasers typically occur in medial position, placed after the modal auxiliary, as in (1b) and (2). If they express special emphasis, they may precede the modal auxiliary. The reinforcing effect of the co-occurrence of modal auxiliaries and modal adverbs, in particular emphasers, in scientific texts will be the main focus of our investigation.

### 3. Data collection and processing

The data were collected from COCA, which is currently the largest corpus of English texts of mixed genres. The size of COCA is more than 560 million words (20 million words each year 1990–2017).<sup>1</sup> The main advantage of COCA is that it includes subdivisions of academic specialized texts and thus represents a good source of scholarly texts. The subcorpus of academic texts is divided into ten subdomains. For the purposes of our analysis, we selected the two subdomains of medicine (ACAD: Medicine) and humanities (ACAD: Humanities). ACAD: Medicine includes 6.7 million words. ACAD: Humanities is larger, it contains 11.9 million words.

The function *Collocate* was used to search for the modal auxiliaries in combination with five modal adverbs in medicine and humanities. The maximal distance from a modal auxiliary on both sides was 3. The selection of modal adverbs was based on Hoyer (1997: 157–158). The five modal adverbs *actually, certainly, clearly, of course, and surely* are frequently used as emphasers.<sup>2</sup> This

also means that they play an important role in the representation of argumentation in research papers. The total number of occurrences of the combinations of MODAL AUXILIARY + MODAL ADVERB (emphasizer) was 1235. Out of these, 968 combinations of modal markers were extracted from ACAD: Humanities and 267 combinations from ACAD: Medicine.

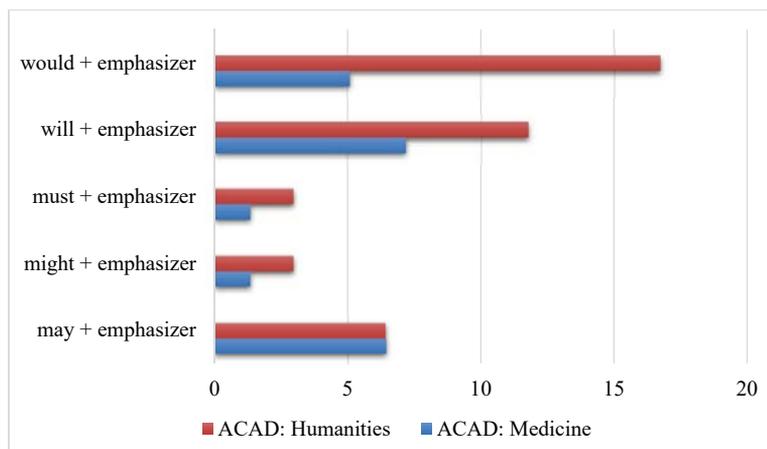
In the next step, all the combinations were manually evaluated. First, only the combinations of MODAL AUXILIARY + MODAL ADVERB (emphasizer) which occurred in one clause were included. Modal meanings expressed by modal auxiliaries were determined. Combinations with a non-epistemic modal meaning were excluded. This resulted in a final set of 440 combinations, 337 in ACAD: Humanities and of 103 in ACAD: Medicine. The results were statistically processed by the log-likelihood test (LL-test) in order to test statistical significance.

The LL-test is preferred by a number of authors (e.g. McEnery, Xiao and Tono 2006: 55) because it “does not assume the data are normally distributed”. In order to determine a measure of likelihood termed the significance or the  $p$  value, another value *degree of freedom* (d.f.) is necessary. It is calculated by multiplying the number of rows less 1 with the number of columns less 1 in a frequency (contingency) table. The LL critical values with 1 d.f. are 3.84 ( $p < 0.05$ ), 95 percentile; 6.63 ( $p < 0.01$ ), 99 percentile; 10.83 ( $p < 0.001$ ), 99.9 percentile; and 15.13 ( $p < 0.0001$ ), 99.99 percentile. McEnery, Xiao and Tono (2006: 56) point out that there are many web-based LL calculators. In addition, the LL-test is a part of standard statistical packages, such as the statistical package for social sciences SPSS. We used the LL calculator by Rayson at Lancaster University available online.<sup>3</sup> The calculator statistically compares the frequency of a word or a phrase, in our analysis a combination MODAL AUXILIARY + MODAL ADVERB (emphasizer), in two subdomains of COCA to test whether an observed difference arises merely due to chance or it reflects a significant link between the two corpora. The significance level set for this study was  $p < 0.01$  with a critical value 6.63.

#### 4. Distribution of epistemic modal markers in ACAD: Medicine and ACAD: Humanities

The analysis of epistemic modal markers expressed by a combination of MODAL AUXILIARY + MODAL ADVERB (emphasizer) resulted in several interesting observations. Figure 2 gives an overview of the distribution of modal combinations in the domain of medicine and humanities.

The frequency values in Figure 2 were normalized per one million words. The results in Figure 2 clearly show that the combinations of modal auxiliaries with the selected emphasizees are much more frequent in the academic domain of humanities. The only exception is represented by the combinations of *may*, which displays very similar normalized frequencies. Hyland (1998), Fløttum et al. (2008), and Ngula (2017) consider *may* a typical marker of epistemic modality in academic texts. Epistemic modal use of combinations with *must* and *might* is



**Figure 2.** Epistemic modal combinations in two academic subdomains in COCA per million words

higher in humanities. However, the calculation of LL-values reveals an important difference. For the modal combinations with *must*, the LL-value 19.15 is statistically significant at the level  $p < 0.01$  whereas for *might*, the LL-value 5.06 is below the critical value 6.63.

Combinations with *would* and *will* are more prevalent in scholarly texts in the humanities. These combinations are often used to hypothesize and present possible interpretations of the data even if the numbers are modest. The examples of combinations with *would* express less certain claims especially when compared to the epistemic meaning of the combinations with *will*. LL-values of 9.47 for the former and 53.53 for the latter significantly deviate from chance (99 percentile). This finding partially corresponds with Ngula (2017: 17–18) where epistemic *would* and *will* ranked among the top five modal verbs in the corpus of research papers in sociology, law and economics.

Figure 2 does not include very low frequencies of combinations with *could*. These were found only in the subcorpus of medicine with an absolute frequency of less than 5. The combinations with *can* were excluded in data processing at the level of semantic analysis. In line with Coates (1983: 19), we found that *can* used in its positive form never expresses epistemic meaning. In negative form, it is used epistemically as a negative of *must*. The reason is that the negative form *must not* expresses exclusively non-epistemic meaning. In our sample, no occurrences of epistemic *should* were recorded either.

#### 4.1. Epistemic modal combinations *WOULD* + *EMPHASIZER*

Combinations of epistemic *would* with the emphasers *actually*, *certainly*, *clearly*, *of course*, and *surely* were more frequent in humanities than in medicine (see

Figure 2). Table 1 summarizes significance values for individual combinations compared between the subdomain of humanities and medicine.

**Table 1.** LL-values for epistemic combinations of WOULD + EMPHASIZER in ACAD: Humanities and ACAD: Medicine

	Normalized f. per million words ACAD: Humanities	Normalized f. per million words ACAD: Medicine	LL-value
<b>would</b> + <b>surely</b>	4.11	0.14	36.1
<b>would</b> + <b>certainly</b>	8.23	1.94	33.91
<b>would</b> + <b>of course</b>	2.68	0.29	17.46
<b>would</b> + <b>clearly</b>	0.50	1.34	3.55
<b>would</b> + <b>actually</b>	1.17	1.34	0.10

The frequencies per one million words in Table 1 show that the combinations with *clearly* and *actually* are more frequent in medicine but the difference is not statistically significant. In contrast, the combinations with *surely*, *certainly*, and *of course* occur more often in humanities. The LL-test carried out to compare the modal combinations in humanities and medicine returned significant differences for these combinations. The data in Table 1 were sorted by the last column with the LL-values in order to highlight significant results.

When scholars use *would* in its epistemic sense, they generally demonstrate less confidence about the claim they make when compared to *will*. Both modal auxiliaries *would* and *will* “exhibit a marked tendency to co-occur with adverbs connoting various degrees of probability (*probably*, *presumably*, *undoubtedly*) and thus express epistemic modality” (Hoye 1997: 114). Some examples from our sample are given in (4).

- (4) a. Such multiple interventions **would** almost **certainly** require substantial financial investments. [ACAD: Medicine]  
 b. First, they **would of course** have to be dedicated to the profession. [ACAD: Medicine]  
 c. Any Mexican politician interested in gaining the attention of the populace **would surely** choose and have access to an electronic medium. [ACAD: Humanities]

In (4a) we can see that the writer is closer to the doubtful end of the continuum when using *would*. On the other hand, the combination with *certainly* adds slightly more emphasis and moves the meaning a bit closer towards confidence at the other end of the continuum. In addition, there is another adverb *almost* in between modal markers with downtoning affect. A similar effect can be seen in (4b). What is perhaps possible is emphasized by the modal adverb. In (4c) the author expresses his confidence in the truth of the proposition, but does not make it explicit that what he says is true. The combination with *surely* puts more emphasis on the truth value.

#### 4.2. Epistemic modal combinations WILL + EMPHASIZER

Epistemic modal combinations of *will* plus *actually*, *certainly*, *clearly*, *of course*, and *surely* turned out to be more frequent in ACAD: Medicine. Figure 2 showed that these combinations are also important in ACAD: Humanities. The frequency per million words for each combination together with statistical significance determined by LL-value calculations is presented in Table 2.

**Table 2.** LL-values for epistemic combinations of WILL + EMPHASIZER in ACAD: Humanities and ACAD: Medicine

	Normalized f. per million words ACAD: Humanities	Normalized f. per million words ACAD: Medicine	LL-value
<b>will + surely</b>	3.61	0.44	22.36
<b>will + of course</b>	2.18	1.04	3.41
<b>will + actually</b>	0.92	1.94	3.27
<b>will + certainly</b>	4.20	2.68	2.82
<b>will + clearly</b>	0.84	1.04	0.19

Table 2 demonstrates that only the two epistemic modal combinations *will + actually* and *will + clearly* occur more often in medical research papers than in the humanities. The combination *will + surely* displays a statistically significant LL-value (22.36), which demonstrates that the difference in distribution between the two subdomains deviates significantly from chance (99 percentile).

Hoye (1997: 117–118) and Coates (1983: 183) note that there is a tendency for epistemic *will* to co-occur with a range of adverbs. They express certainty but they can also express less confidence. Ngula (2017: 11) classifies the modal auxiliary *will* as the one with a strong epistemic force or commitment. His results show that international writers in three humanities disciplines (law, sociology, and economics) are inclined to use the more tentative form *would* to express epistemic modality compared with the stronger form *will*. This is in line with our findings for the subcorpus of research papers in humanities in Table 1, but it contrasts with our data for *will* + *emphasizer* in medicine in Figure 2. Some examples are in (5).

- (5) a. Yet the opportunities for advances in knowledge and the practical application of these advances **will surely** increase. [ACAD: Medicine]  
 b. Initiation of dialogue with such marginal populations **will certainly** help in expanding the outreach of government health care services. [ACAD: Medicine]

Both examples in (5) were collected from ACAD: Medicine. Modal adverbs *surely* in (5a) and *certainly* in (5b) “represent the highest point of the scale of likelihood; as markers of the speaker’s confidence” (Hoye 1997: 119). In (5a), the writer is confident about making the prediction that the practical use of advances he mentions will increase. Similarly, the combination of *will* plus *certainly* in (5b) strengthens the impression that the author of the research paper believes in the truth of the proposition.

#### 4.3. Epistemic modal combinations *MUST* + *EMPHASIZER*

The modal auxiliary *must* is less used to express epistemic meaning (Coates 1983; Hoye 1997). On the other hand, Hoye (1997: 105) gives a list of adverbs that tend to collocate with epistemic *must* such as *apparently*, *obviously*, *clearly*, *inevitably*, and *evidently*. The results for the combinations with *must* + *emphasizer* are in Table 3.

Table 3 shows that the combinations of *must* with *emphasizers* in our sample were more frequent in humanities. No records of the combination *must* + *certainly*, *must* + *actually*, and *must* + *clearly* were found in ACAD: Medicine. The analysis by Ngula (2017: 15–16) revealed that *must* was not frequent enough in research papers either by international researchers or Ghanaian non-native speakers of English. As a result, it did not appear among his top five modal verbs in three disciplines in the humanities. In Table 3 there are two combinations which are statistically significant in the comparison of frequencies between humanities and medicine. These are *must* + *certainly* and *must* + *surely*. Some examples of epistemic uses of *must* in combinations are in (6).

**Table 3.** LL-values for epistemic combinations of MUST + EMPHASIZER in ACAD: Humanities and ACAD: Medicine

	Normalized f. per million words ACAD: Humanities	Normalized f. per million words ACAD: Medicine	LL-value
<b>must</b> + <b>certainly</b>	1.17	0.00	12.51
<b>must</b> + <b>surely</b>	2.52	0.59	10.34
<b>must</b> + <b>actually</b>	0.42	0.00	4.47
<b>must</b> + <b>clearly</b>	0.33	0.00	3.57
<b>must</b> + <b>of course</b>	0.67	0.74	0.03

- (6) a. “Key criteria” for sending a patient to a colorectal clinic **must surely** depend on the importance of the symptoms. [ACAD: Medicine]  
 b. While we **must certainly** be cautious of a seemingly too-easy structural analysis, we should not avoid it either--if it aids us in understanding African cultures. [ACAD: Humanities]

Combinations of epistemic *must* + *adverb* often reflect the speaker’s degree of confidence or inferential nature of *must* (see Figure 1). In (6a) we deal with the epistemic use of *must* and it can be paraphrased as that the speaker confidently infers that there must be a typology of symptoms which then serves as a basis for making a decision whether it is necessary to send a patient to a colorectal clinic. This meaning has been labelled as *epistemic necessity* (Coates 1983; Hoye 1997). In (6b) the writer is certain about being cautious and it is very likely that this viewpoint is supported by evidence.

#### 4.4. Epistemic modal combinations MIGHT + EMPHASIZER

In line with Hoye (1997: 96), *might* together with *may* primarily expresses epistemic possibility and in combinations with adverbs, they often relate to the higher value of probability. The research study by Hoye (1997: 94) demonstrated that *might* + *adverbial* expressed epistemic modal meaning in 92 per cent. The results of our analysis are in Table 4.

**Table 4.** LL-values for epistemic combinations of MIGHT + EMPHASIZER in ACAD: Humanities and ACAD: Medicine

	Normalized f. per million words ACAD: Humanities	Normalized f. per million words ACAD: Medicine	LL-value
<b>might + actually</b>	2.10	1.04	3.01
<b>might + of course</b>	0.67	0.14	2.91
<b>might + clearly</b>	0.00	0.14	2.04
<b>might + certainly</b>	0.16	0.00	1.79
<b>might + surely</b>	0.00	0.00	0.00

The values in Table 4 clearly demonstrate that the combinations of *might* + emphasizer are very rare in scientific text in medicine and humanities. No LL-value is statistically significant for the difference between the two subcorpora. This result is different from Ngula (2017: 15-16) who found that *might* ranks among the top five modal verbs used in research articles not only in Ghanaian authors but also by English scholars.

According to Høye (1997: 97) *might* and *may* are “subject to intensification where the effect of the adverb modifier can be to diminish or weaken the force of the modal”. Some examples from our set are given in (7).

- (7) a. Immune responses restored by HAART **might actually** exacerbate progression of HBV infection, as has been reported for tuberculosis and cytomegalovirus infections. [ACAD: Medicine]  
 b. What decisions and actions **might actually** make it more difficult for students to succeed? [ACAD: Humanities]

In (7a) and (7b) we can see that the writers consider that X stated in each sentence is perhaps possible. This means that these combinations are rather weak in their force. The frequencies for these modal combinations are surprisingly low, which contrasts sharply with Ngula’s observations for *might*, ranking it among the top five modals used in academic papers.

#### 4.5. Epistemic modal combinations *MAY* + *EMPHASIZER*

Hoye (1997: 94) demonstrated that *may* + ADVERBIAL expressed epistemic modal meaning in 76 per cent. The values for the combinations of *may* + EMPHASIZER are in Table 5.

**Table 5.** LL-values for epistemic combinations of *MAY* + *EMPHASIZER* in ACAD: Humanities and ACAD: Medicine

	Normalized f. per million words ACAD: Humanities	Normalized f. per million words ACAD: Medicine	LL-value
<b>may + of course</b>	1.76	0.74	3.51
<b>may + certainly</b>	0.25	0.00	2.68
<b>may + clearly</b>	0.25	0.00	2.68
<b>may + actually</b>	4.03	5.67	2.42
<b>may + surely</b>	0.08	0.00	0.89

The data in Table 5 show that the frequencies of epistemic combinations of *MAY* + *emphasizer* are very low. The LL-test did not prove any statistical significance at the set level of  $p < 0.01$  with a critical value of 6.63. Examples of the epistemic meaning of these combinations are given in (8).

- (8) a. However, propositions **may actually** form a bridge between linguistic and conceptual metaphor and have proven to be a very useful tool in our approach. [ACAD: Humanities]  
 b. Such figures indicate that the degree of risk incurred by the youth **may actually** be amplified once they begin living on the streets. [ACAD: Medicine]

The examples in (8) express epistemic possibility. The author of (8a) indicates that the speaker thinks what is perhaps possible. In combination with *actually*, it relates to the higher degree of probability. A similar epistemic meaning appears in (8b). The author is not one hundred per cent confident but considers the claim probable.

## 5. Conclusion

This study investigated the distribution of epistemic modal combinations of MODAL AUXILIARY + MODAL ADVERB (emphasizer) in the corpus of medical research papers ACAD: Medicine and research papers in humanities ACAD: Humanities in COCA (2008-2018). The analysis showed that there are statistically significant differences between medical research papers and articles from the humanities, especially in the use of epistemic modal combinations with *would*, *must*, and *will*. Scholars in the humanities tend to use epistemic modal combinations more frequently than their medical counterparts.

Statistically significant differences between these two domains were observed in the use of epistemic modal combination of *would* + *surely*, *would* + *certainly*, and *would* + *of course*. For the combinations with *must*, two combinations are significant, *must* + *certainly* and *must surely*. The combinations with *will* revealed only a single combination which is statistically significant i.e. *will* + *surely*. The statistical figures show that in general, epistemic modal markers and their combinations are more frequent in humanities than in medicine.

These findings suggest that there indeed are differences in argumentation line of research papers in different scientific domains. The main role of epistemic modal markers is to assess the chances or the likelihood of claims presented by the authors. Combinations of epistemic modal markers are indispensable especially in evaluating how confident writers are in presenting their results and claims.

One of the potential reasons for the differences of preferred patterns could be the nature of the disciplines themselves and the difference in their aim. In science, represented by the subcorpus of medicine, the main aim is assumed to be to inform the reader about the results and implications of experiments or the interpretation of quantitative data in an objective manner. On the other hand, in the humanities, authors often do not rely on hard data but rather on an interpretation of the source material. The interpretation is conveyed to the reader through argumentation and persuasion, giving the author more freedom in making stronger claims that directly follow from quantitative data. The results also indicate that in the medical domain, sentential adverbial expressions are quite rare because there is a high prevalence of objective, fact-based conclusions being reported in medical texts. Humanities texts, on the other hand, can be much more speculative in nature. Further research, comparing more different domains, is needed to search for more combinations of modal auxiliaries across different disciplines and potentially determining the patterns which are domain-specific.

## Notes

<sup>1</sup> Data available at <http://corpus.byu.edu/coca>, retrieved 7 January 2019.

<sup>2</sup> It should be noted that Hoyer (1997: 157–158) treats *of course* as a single unit. Here we adopt the same approach.

- <sup>3</sup> Created by Paul Rayson at Lancaster University, available at <http://ucrel.lancs.ac.uk/llwizard.html>, retrieved 8 January 2019.

## References

- Aarts, Baas (2011) *Oxford Modern English Grammar*. Oxford: Oxford University Press.
- Ardizzone, Patrizia and Pennisi, Giulia A. (2012) Epistemic modality variation in community law journals. In: Stefania M. Maci and Michele Sala (eds.) *Genre Variation in Academic Communication: Emerging Disciplinary Trends*. Bergamo: CELSB, 153–174.
- Bolinger, Dwight (1972) *Degree Words*. The Hague and Paris: Mouton.
- Boye, Kasper (2012) A Usage-Based Theory of Grammatical Status and Grammaticalization. *Language* 16 (1), 9–43.
- Boye, Kasper (2016) The Expression of Epistemic Modality. In: J. Nuyts and J. Van Der Auwera (eds.) *The Oxford Handbook of Modality and Mood*. Oxford: Oxford University Press, 117–140.
- Bybee, Joan, L., Perkins, Revere D., and Pagliuca, William (1994) *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chovanec, Jan (2012) Written academic discourse in English: From local traditions to global outreach. *Brno Studies in English* 38 (2), 5–16.
- Coates, Jennifer (1983) *The Semantics of the Modal Auxiliaries*. London: Croom Helm.
- Crystal, David (2012) *English as a Global Language*. Cambridge: Cambridge University Press.
- Davies, Mark (2008-) *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available online at <https://corpus.byu.edu/coca/>.
- Fergusson, Gibson (2007) The global spread of English, scientific communication and ESP: questions of equity, access and domain loss. *Ibérica* 13, 7–38.
- Fløttum, Kjersti, Trine Dahl, Torodd Kinn, Anje M. Gjesdal and Eva T. Vold (2008) Cultural identities and academic voices. In: Fløttum, Kjersti (ed.) *Language and Discipline Perspectives on Academic Discourse*. Newcastle upon Tyne: Cambridge Scholars Publishing, 14–39.
- Flowerdew, John (2007) Scholarly writers who use English as an Additional Language: What can Goffman's "Stigma" tell us? *Journal of English for Academic Purposes* 7, 77–86.
- Giannoni, Davide (2008) Medical writing at the periphery: The case of Italian journal editorials. *Journal of English for Academic Purposes* 7, 97–107.
- Gunnarsson, Britt-Louise (2009) *Professional Discourse*. London: Continuum.
- Hermerén, Lars (1986) Modalities in spoken and written English. An inventory of forms. In: Gunnel Tottie and Ingegerd Backlund (eds.) *English in Speech and Writing: A Symposium*. Studia Anglistica Upsaliensia 60. Stockholm: Almqvist & Wiksell, 57–91.
- Holmes, Janet (1983) Speaking English with the appropriate degree of conviction. In: Brumfit, C. (ed.) *Learning and Teaching Languages for Communication: Applied Linguistic Perspectives*. London: Centre for Information on Language Teaching and Research, 100–113.
- Hoye, Leo (1997) *Adverbs and Modality in English*. London and New York: Longman.
- Huddleston, Rodney and Pullum, Geoffrey K. (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- 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 and Philadelphia: John Benjamin Publishing.
- Hyland, Ken (2001) Definitely a possible explanation: epistemic modality in academic argument. In: Maurizio Gotti and Marina Dossena (eds.) *Modality in Specialized Texts*. Bern: Peter Lang, 291–230.
- Hyland, Ken (2012) English for academic purposes and discourse analysis. In: J. P. Gee and M. Handford (eds.) *The Routledge Handbook of Discourse Analysis*. London: Routledge, 412–423.

- Kranich, Svenja and Gast, Volker (2015) Explicitness of epistemic modal marking: recent changes in British and American English. In: J. R. Zamarrano-Mansilla et al. (eds.) *Thinking Modality: English and Contrastive Studies of Modality*. Newcastle upon Tyne: Cambridge Scholars Publishing, 3–22.
- Large, J. Andrew (1983) *The Foreign-Language Barrier: Problems in Scientific Communication*. London: André Deutsch.
- Lyons, John (1977) *Semantics*. Cambridge: Cambridge University Press.
- Maher, John (1986) The development of English as an international language of medicine. *Applied Linguistics* 7, 206–218.
- McEnery, Tony, Xiao, Richard and Tono, Yukio (2006) *Corpus-based Language Studies: An Advanced Resource Book*. London and New York: Routledge.
- Myers, Greg (1990) *Writing Biology: Texts in the Social Construction of Scientific Knowledge*. London: University of Wisconsin Press.
- Ngula, Richmond S. (2015). *Epistemic Modality in Social Science Research Articles Written by Ghanaian Authors: A Corpus-Based Study of Disciplinary and Native vs. Non-Native Variations*. Unpublished PhD thesis, Lancaster University, UK.
- Ngula, Richmond S. (2017) Epistemic modal verbs in research articles written by Ghanaian and international scholars: A corpus-based study of three disciplines. *Brno Studies in English* 43 (2), 5–27.
- Nuyts, Jan (2001) *Epistemic Modality, Language and Conceptualization: A Cognitive Pragmatic Perspective*. Amsterdam and Philadelphia: John Benjamins.
- Nuyts, Jan (2016) Analyses of modal meanings. In: Jan Nuyts and Johan Van Der Auwera (eds.) *The Oxford Handbook of Modality and Mood*. Oxford: Oxford University Press, 31–49.
- Orta, Ignacio Vázquez (2010) A contrastive analysis of the use of modal verbs in the expression of epistemic stance in Business Management research articles in English and Spanish. *Ibérica* 19, 77–96.
- Palmer, Frank (1986) *Mood and Modality*. Cambridge: Cambridge University Press.
- Palmer, Frank (2001) *Mood and Modality*. 2nd ed. Cambridge: Cambridge University Press.
- Panocová, Renáta (2008) Expressions of modality in biomedical texts. *SKASE Journal of Translation and Interpretation* 3 (1), 82–90.
- Pérez-Llantada, Carmen (2010) The ‘dialectics of change’ as a facet of globalisation: The Epistemic modality in academic writing. In: M. F. Ruiz-Garrido, J. C. Palmer, I. Fortanet-Gómez, I. Fortanet (eds.) *English for Academic Purposes*. Amsterdam and New York: Rodopi, 25–42.
- Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey, and Svartvik, Jan (1985) *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- Salager-Meyer, Françoise (1994) Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes* 13 (2), 149–170.
- Skelton, John (1997) How to tell the truth in The British Medical Journal: Patterns of judgement in the 19th and 20th centuries. In: R. Markkanen and H. Schröder (eds.), *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*. Berlin: Walter de Gruyter, 42–58.
- Ševčíková, Magda (2009) *Funkce kondicionálu z hlediska významové roviny*. Praha: Ústav formální a aplikované lingvistiky.
- Thompson, Geoff (2001) Interaction in academic writing: learning to argue with the reader. *Applied Linguistics* 22 (1), 58–78.
- Varttala, Teppo (2003) Hedging in scientific research articles: A cross-disciplinary study. In: Giuseppina Cortese and Phillip Riley (eds.) *Domain-Specific English: Textual Practices across Communities and Classrooms*. New York: Peter Lang, 141–174.
- Verstraete, Jean-Christophe (2001) Subjective and Objective Modality: Interpersonal and Ideational Functions in the English Modal Auxiliary System. *Journal of Pragmatics* 33, 1505–1528.
- Vold, Eva Thue (2006) Epistemic modality markers in research articles: A cross-linguistics and cross-disciplinary study. *International Journal of Applied Linguistics* 16 (1), 62–87.
- Von Wright, Georg H. (1951) *An Essay in Modal Logic*. Amsterdam: North-Holland.

RENÁTA PANOCOVÁ is an associate professor at the Department of British and American Studies of Pavol Jozef Šafárik University in Košice, Slovakia. Her research interests include morphology, word formation, and terminology. She is the author of *Categories of Word Formation and Borrowing: An Onomasiological Account of Neoclassical Formations* (2015), *The Vocabulary of Medical English: A Corpus-based Study* (2017) and co-editor of *Word Formation and Transparency in Medical English* (2015).

Address: Renáta Panocová, Pavol Jozef Šafárik University in Košice, Faculty of Arts, Šrobárova 2, Košice, 040 01, Slovakia. [renata.panocova@upjs.sk]

LUKÁŠ LUKAČÍN is a PhD student at the Department of British and American Studies of Pavol Jozef Šafárik University in Košice, Slovakia. His PhD thesis focuses on the investigation of modal markers in medical English.

Address: Lukáš Lukačín, Pavol Jozef Šafárik University in Košice, Faculty of Arts, Šrobárova 2, Košice, 040 01, Slovakia. [lukas.lukacin@student.upjs.sk]



This work can be used in accordance with the Creative Commons BY-NC-ND 4.0 International license terms and conditions (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>). This does not apply to works or elements (such as image or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.