

May and *must* as inferential evidential markers in a corpus of English scientific texts

Ivalla ORTEGA BARRERA and Elena QUINTANA TOLEDO¹

Universidad de Las Palmas de Gran Canaria
Departamento de Filología Inglesa
iortega@dfm.ulpgc

Recibido: 16/10/2012

Aceptado: 23/11/2012

ABSTRACT

This paper explores the use of *may* and *must* as inferential evidential markers in the *Corpus of Specialized Papers in English*, currently in progress at the Institute for Technological Development and Innovation in Communications at the University of Las Palmas de Gran Canaria. In order to do so, we will follow Cornillie's (2009) disjunctive approach to the concepts of evidentiality and epistemic modality to firstly identify evidential rather than epistemic readings of these modals. Their qualitative assessment as inferential evidential markers will be carried out through the analysis of the context in which they are immersed. We will show that appropriate contextual enrichment allows the recovery of the intended meaning, that is, mode of knowing, of both *may* and *must*.

Key words: Evidentiality, epistemic modality, context, modals, disjunctive approach

May y *must* como marcadores evidenciales inferenciales en un corpus de textos científicos en inglés

RESUMEN

En este artículo se examina el uso de *may* y *must* como marcadores de evidencialidad inferencial en el *Corpus de Artículos Especializados en Inglés*, que se desarrolla actualmente en el Instituto para el Desarrollo Tecnológico y la Innovación en las Comunicaciones perteneciente a la Universidad de Las Palmas de Gran Canaria. Para realizar este análisis secundaremos el enfoque disyuntivo de Cornillie (2009) en el estudio de los conceptos de evidencialidad y modalidad epistémica para identificar la naturaleza evidencial y no epistémica de estos modales. Su evaluación cualitativa como marcadores de evidencialidad inferencial se realizará a través del análisis del contexto en el que se encuentran inmersos. Finalmente mostraremos que un enriquecimiento contextual apropiado permite identificar que el uso de estos dos modales marcan el modo de conocimiento de los autores.

Palabras clave: Evidencialidad, modalidad epistémica, contexto, modales, enfoque disyuntivo

¹ D. Ivalla Ortega-Barrera is part of the research project *Evidencialidad en un corpus multidisciplinario de artículos científico-técnicos en lengua inglesa*, grant FFI2009-10801 (FEDER, Spanish Ministry for Science and Innovation). Elena Quintana-Toledo collaborates in the same research project. This grant is hereby gratefully acknowledged.

SUMMARY: 1. Introduction. 2. Corpus description. 3. Evidentiality and modality. 3.1. Epistemic and deontic *may* and *must*. 3.2. Classification of evidential values. 4. Analysis of *may* and *must* as inferential evidential markers. 4.1. Inferential evidential *must*. 4.2. Inferential evidential *may*. 5. Conclusion.

1. INTRODUCTION

The main purpose of this paper is the study of the use of the modal verbs *may* and *must* as inferential evidential markers in a corpus of English scientific texts. We will show that both modal verbs are used as inferential evidential markers in the introduction and conclusion sections of English research papers by examining the context in which they occur.

In the modal logic tradition, the semantics of *may* expresses dynamic possibility, deontic permission and epistemic possibility (Palmer 1990). Van der Auwera refers to *may* as an indicator of indeterminacy since it expresses what is “neither T[rue] nor F[false] but also either T[rue] or F[false]” (Van der Auwera 1986: 1068). As for *must*, some of the meanings attached to it are logical conclusion, necessity and obligation (Klinge 1993: 351).

Evidentiality is a functional category that can be understood in a narrow sense as the linguistic coding of the speaker’s source of knowledge (Givón 1982; Bybee 1985; Chafe 1986; Mithun 1986; Willet 1988; Mayer 1990). In a broad sense, it does not only refer to the speaker’s source of knowledge but also to the degree of certainty as for the proposition expressed (Givón 1982; Chafe 1986; Palmer 1986; Mayer 1990). This broad concept has led to the association of the notions of evidentiality and epistemic modality. The relationship between the two may be approached from different perspectives: some scholars base their research on the idea that evidentiality and epistemic modality are distinct categories (De Haan 1999; Nuyts 2001; Aikhenvald 2004; Cornillie 2009). Others take one of the two notions to be embedded within the other, that is, evidentiality within epistemic modality (Palmer 1986; Willet 1988), or epistemic modality within evidentiality (Papafragou 2000; Ifantidou 2001). Another angle from which this issue can be discussed is that of overlap (Van der Auwera and Plungian 1998; Palmer 2001).

The present paper will follow a disjunctive view as proposed by Cornillie (2009) and so we will consider them as conceptually different categories. Even though this view does not exclude the possibility that modals may present both evidential and epistemic readings, we will concentrate on the former to show that the modal verbs *may* and *must* are indicators of mode of knowing and that knowledge has been acquired inferentially.

The paper is organised as follows: Section 2 offers the description of the corpus of study. Section 3 contains a brief account of evidentiality with special focus on its relationship with modality. Section 4 deals with the analysis and discussion of *may*

and *must* as inferential evidential markers. The last section presents the findings and conclusions drawn from the present study.

2. CORPUS DESCRIPTION

The data analysed to carry out this study has been excerpted from the *Corpus of Specialized Papers in English*. This corpus is part of a larger research project, *Evidentiality in a Multidisciplinary Corpus of Research Papers in English*, currently in progress at the Institute for Technological Development and Innovation in Communications at the University of Las Palmas de Gran Canaria.

The corpus contains research papers from 1998 to 2008 and covers three register domains, namely computing, law and medicine. They have been randomly selected from several databases of scientific journals, which include *ACM Computing Surveys*, *IEEE Transactions on Graphics*, *International Journal of Computer Vision*, *Air & Space Law*, *American Journal of Criminal Law*, *Business Law Review*, *Academic Emergency Medicine*, *Aesthetic Plastic Surgery* and *Basic Applied Myology*, among others.

In an attempt to analyse the use of *may* and *must* as inferential evidential markers, we have selected thirty papers per register domain (law, medicine and computing) belonging to journals with a high impact factor from the *Corpus of Specialized Papers in English* (they can be found in the appendix). Thus, we have taken into account sociological characteristics to select the articles for our study; for instance, these have to be written in English and at least one of their authors must be a native speaker of the language. All these features would allow for a unified account of the findings.

To carry out this study, we have focused on the introductory and concluding sections because it is in these sections where we find different rhetorical functions. In the first one, the authors present their research, give the theoretical framework, define the study and justify it; in the second one, the authors justify and stress the importance of their first statement; they explain it, discuss it and generalise it.

Introductions and conclusions in research articles have been assessed in terms of moves, that is, the parts which accomplish specific rhetorical functions within the whole structure of the section. Swales' (1990: 141) *Create A Research Space* (CARS) model, outlined in Table 1, has been the most widely used for the assessment of introductory sections. It has been taken as the starting point in the analysis of introductions of research articles belonging to varied disciplines such as electronic engineering (Cooper 1985), social science (Crookes 1986), medicine (Ngozi-Nwogu 1997), software engineering (Anthony 1999), computer engineering (Posteguillo 1999) or environmental science (Samraj 2002), to mention just a few.

Although these works evince that the structural organisation of introductions show disciplinary variation, the pattern identified in Swales' model "is frequently found in

more or less its pure form in many disciplines” (Dudley-Evans 2000: 6). In Table 1 we present the moves in the introductions of research articles according to the CARS model:

Move	
Establishing a territory	Step 1 Claiming centrality and/or Step 2 Making topic generalization(s) and/or Step 3 Reviewing items of previous research
Establishing a niche	Step 1A Counter claiming or Step 1B Indicating a gap or Step 1C Question-raising or Step 1D Continuing a tradition
Occupying the niche	Step 1A Outlining purposes or Step 1B Announcing present research Step 2 Announcing principal findings Step 3 Indicating RA structure

Table 1. Moves in the introductions of research articles according to the CARS model (Swales 1990: 141).

As for conclusions, these are not always treated as independent sections within the whole structure of the research article, but as embedded within the discussion section where they occupy a closing position (Dudley-Evans 1986, 1994). We have encountered this in some of the articles analysed, mainly in those belonging to the medical register, so in these cases we have carefully read the discussion sections so as to identify those paragraphs aimed at providing some sort of conclusion, i.e. summary of results and claims as well as recommendations for future research.

Computing and legal texts, in contrast, tend to separate the discussion and conclusion sections. Some research has been conducted on the independent status of conclusions, for instance in Ruiying and Allison (2003), whose proposal for the structure of this stage in applied linguistics articles is shown in Table 2. Texts in our corpus generally present the moves and steps identified by these authors, except for the drawing of pedagogic implications.

Move	
Summarising the study	
Evaluating the study	Indicating significance/advantage Indicating limitations Evaluating methodology
Deductions from research	Recommending further research Drawing pedagogic implication

Table 2. Moves in the conclusions of research articles (Ruiying & Allison 2003: 379).

All in all, variation across disciplines in the structural organisation of research articles does not seem to be the result of authorial choice. It is importantly determined by the discipline itself as well as by the guidelines imposed by the editorial committee of the journal they are published in.

The methodology followed to accomplish this study combines computer analysis for the initial detection of the modals under survey with manual analysis for the identification of their inferential evidential value. As already stated, we have also read carefully the discussion sections of medical texts to identify those paragraphs where authors indicate potential conclusions for the paper. We have used the *Online Interface for Corpus Management* (OnICoMt) in order to make the computerised searches. OnICoMt is an information retrieval tool that the research group *Tecnologías Emergentes aplicadas a la Lengua y la Literatura* at the University of Las Palmas de Gran Canaria has designed and implemented to accompany the corpus.

3. EVIDENTIALITY AND MODALITY

Evidentiality may be generally defined as the linguistic coding of the speaker's source of knowledge. Formal treatments of evidentiality tend to associate it with epistemic modality, even though they are conceptually different (Dendale and Tasmowski 2001; McCready and Ogata 2007; Cornillie 2009), since the former refers to mode of knowing and the latter is related to the degree of speaker's commitment towards propositional content. The relationship between evidentiality and epistemic modality has been addressed from three different perspectives, namely, inclusion, overlap and disjunction (Dendale and Tasmowski 2001).

The inclusive approach takes either of the two notions as a subdomain of the other, and so the term evidentiality does not only refer to the speaker's source of knowledge but also to the degree of reliability of that knowledge. Palmer (1986) endorses the view that evidentiality is subsumed under the epistemic modal system and considers evidential markers as indicators of the speaker's commitment to the truth of the proposition expressed. He defines the notion of epistemicity as follows:

the term 'epistemic' should apply not simply to modal systems that basically involve the notions of possibility and necessity, but to any modal that indicates the degree of commitment by the speaker to what he says. In particular, it should include evidentials such as 'hearsay' or 'report' (the quotative) or the evidence of the senses (Palmer 1986: 51).

On the same line, Aijmer (1980) associates epistemicity with the expression of both evidence and certainty in the following terms: "Epistemic quantifiers are expressions which say something about the speaker's evidence and degree of certainty" (Aijmer 1980: 11). Dendale and Tasmowski (2001) point out that inclusive

approaches of this type are often motivated by the idea that indicating source of knowledge is an indirect way of conveying epistemic attitude.

Van der Auwera and Plungian (1998), on their part, advocate for an overlapping approach. They take epistemic modality as “a domain where evidential and modal values overlap [...] where the probability of P is evaluated” (Plungian 2001: 354). Their definition of evidentiality is related to “the indication of the source or kind of evidence speakers have for their statements” (Van der Auwera and Plungian 1998: 85). Relying on the semantic mapping theory, they consider that the category of epistemic modality includes two meaning regions: (i) epistemic possibility, i.e. uncertainty, and (ii) epistemic necessity, i.e. certainty and a relatively high degree of probability. The meaning regions covered by the evidential category are direct, i.e. visual, auditory and unspecified, and indirect, i.e. reportatives, inferential and unspecified. Their proposal equates inferential evidential readings with epistemic necessity (Van der Auwera and Plungian 1998: 85; cf. Boye 2010).

Carretero (2004) also deals with the relationship between evidentiality and epistemic modality from an overlapping perspective. She emphasises the existence of a continuum from evidentiality to epistemicity so that a given expression will fall towards one point or another of the continuum “depending on the commitment to the truth of the utterance which they encode or implicate” (Carretero 2004: 27-28).

Finally, Cornillie (2009) adopts a disjunctive standpoint in which evidentiality and epistemic modality are considered as entirely distinct categories. He posits that the expression of mode of knowing does not necessarily correlate with the expression of epistemic speaker commitment. Drawing on Nuyts’ definition of epistemic modality, that is, “an evaluation of the chances that a certain hypothetical state of affairs under consideration (or some aspect of it) will occur, is occurring or has occurred in a possible world” (Nuyts 2001: 21), he suggests taking evidentiality as “the reasoning processes that lead to a proposition” and epistemic modality as the evaluation of “the likelihood that this proposition is true” (Cornillie 2009: 46-47). This framework allows for the occurrence of multiple readings of a single expression which, as we will see later, happens to be frequent in the case of modals.

As far as modals are concerned, Palmer (1986: 21-23) defines modality in contrast to mood as a linguistic phenomenon associated with the use of several linguistic features among which modals occupy a prominent place. Mood, however, is a category that is typically morphologically realised. Traditionally, modals may encode epistemic or deontic meanings. The former are expressed by those linguistic items that indicate the degree of speaker’s certainty in the truth of the underlying proposition and the latter are expressed by those linguistic items that indicate obligation and permission (De Haan 2006).

Aikhenvald (2004) notes that modals cannot always be easily assigned to the realm of evidentiality or modality in their purest sense; sometimes it is difficult to decide whether a given modal expresses assumption or inference, or epistemicity (Aikhenvald 2004: 150-151). In any case, context is a determining factor for the

identification of evidential markers and so we will address the interpretation of inferential evidential uses of *may* and *must* by studying the context in which they appear and its possible influence in the modal verb.

3.1. EPISTEMIC AND DEONTIC *MAY* AND *MUST*

Modality is concerned with the codification of the speaker's attitude towards "the proposition that the sentence expresses" (Lyons 1977: 452) or, in other words, with the "speaker's assessment of, or attitude towards, the potentiality of a state of affairs" (Radden and Dirven 2007: 234). Modal verbs may express more than one type of modality. As will be apparent from the analysis in section 4 below, *may* and *must* may express evidential meanings. However, these modals may also express other types of modal meanings, mainly epistemic and deontic. Let us consider (1) to (3) for illustration:

- (1) I may be wrong... I don't know
- (2) You may leave now
- (3) You must be in bed by ten o'clock

Example (1) contains an epistemic use of *may*, and so the modal stands as an indicator of the speaker's assessment of the probability of the proposition in terms of possibility. His/her uncertainty in the actualisation of the proposition is further reinforced by the occurrence of *I don't know* which emphasises the speaker's lack of confidence in its truth. Examples (2) and (3), on their part, present deontic uses of the modals. Deontic modality is involved when the speaker "intervene[s] in the speech event by laying obligations or giving permission" (Downing and Locke 1992: 382). The deontic source coincides with the speaker in both cases: in (2) the speaker is granting the hearer permission to leave the place where they both are; in (3) the speaker is laying an obligation upon the hearer.

3.2. CLASSIFICATION OF EVIDENTIAL VALUES

Evidential markers have been traditionally classified into direct and indirect depending on the specific way the speaker has been led to some knowledge; for instance, s/he may have witnessed a given event, s/he may have been told by someone else or s/he may have deduced it on the basis of some evidence at her/his disposal. Aikhenvald divides evidentials in six categories, namely, (i) visual, (ii) non-visual, (iii) inferential, (iv) assumptive, (v) hearsay and (vi) quotative (Aikhenvald 2004: 63-64). Visual and non-visual evidential markers are those which indicate that the information has been acquired through the senses, mostly seeing and hearing, but also

smell, taste and touch. Inferential evidentials point at the information being acquired on the basis of deduction made through evidence (visible or tangible), or result. Alonso-Almeida and Lareo (forthcoming), commenting on Van der Auwera and Plungian (1998: 85-86), write that they “describe inferences as the result of an event in which strong epistemic necessity allows deductions in the light of certain evidences”. Assumptive evidentials show that information has been acquired through general knowledge and assumptions as well as logical reasoning. Hearsay and quotative evidentials are used when the information has been reported with no reference and with reference to those reporting it, respectively.

Aikhenvald’s classification does not really seem to be satisfactory because of the inclusion of logical reasoning within the assumptive category: any process of logical reasoning relies on the individual’s inferential abilities at some point or another in which case information acquired by logical reasoning should be considered within the inferential evidential category. Plungian’s (2001) typology of evidentials, however, seems to provide a clearer classification of evidential values. Partially following Willett’s (1988) classification, he proposes a basic distinction between direct evidence and indirect evidence. While the former can be visual, sensoric and endophoric, the latter can be reflected and mediated. They all mark the relationship between some knowledge and the subject as follows:

Direct evidence (including direct access to P)

Visual: “P, and I see/saw P”.

Sensoric: “P, and I perceive(d) P” [P may be heard, smelled, tasted, etc.].

Endophoric: “P, and I feel (felt) P” [P is the speaker’s inner state, cf. *I am hungry, I want to sleep, I know the answer*, etc.].

Reflected evidence (including direct access to some situation Q related to P)

Synchronous inference: “P, because I can observe some signs of P” [P at T0]; cf. *He must be hungry* (because he shows signs of it, etc.).

Retrospective inference: “P, because I can observe some traces of P” [P before T0]; cf. *He must have slept there* (because we see his untidy bed, etc.).

Reasoning: “P, because I know Q, and I know that Q entails P”; cf. *Today there must be a fair in Salzburg* (because I know the routines of this region, etc.).

Mediated evidence (including neither direct nor reflected access to P)

Quotative: “P, because I was told that P”; cf. *They said he’s leaving; He is said to have left*, etc. (Plungian 2001: 354).

4. ANALYSIS OF *MAY* AND *MUST* AS INFERENTIAL EVIDENTIAL MARKERS

In this section we will focus on the methodology followed to perform our analysis and on the research itself. We have firstly made an electronic search of the items *may* and *must* using OnICoMt so as to detect the occurrence of all instances. We have filtered out results by genre sections in order to retrieve those instances present in the introductory and concluding sections of the research papers. We have added the data from the concluding paragraphs of the discussion sections of the medical texts. The results have been tagged manually as epistemic, deontic or evidential. The total amount of words analysed is c. 50000, but data have been normalised per ten thousand words because of the varying lengths of texts, that is, since not all the texts have the same total amount of words, we have divided the results obtained per ten thousand words to obtain a normal distribution so as to make results comparable. Figure 1 and tables 3 and 4 show the distribution of the different interpretations of *may* and *must* in the introduction and conclusion sections of the scientific papers in our corpus:

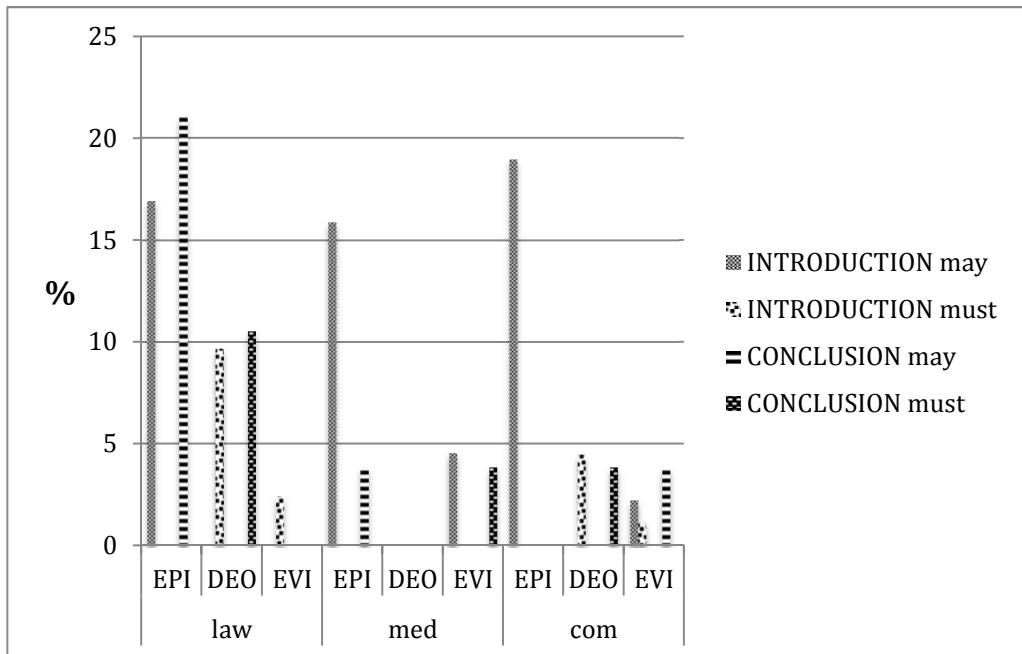


Figure 1. Distribution of interpretations of *may* and *must* in the introductory and concluding sections.

		law			med			com		
		EPI	DEO	EVI	EPI	DEO	EVI	EPI	DEO	EVI
INTRODUCTION	may	16.9	0	0	15.9	0	4.5	18.9	0	2.2
	must	0	9.7	2.4	0	0	0	0	4.4	1.1
		law			med			com		
		EPI	DEO	EVI	EPI	DEO	EVI	EPI	DEO	EVI
CONCLUSION	may	21	0	0	3.8	0	0	0	0	3.8
	must	0	10.5	0	0	0	3.8	0	3.8	0

Table 3. Raw frequencies of *may* and *must* in the corpus analysed

		law			med			com		
		EPI	DEO	EVI	EPI	DEO	EVI	EPI	DEO	EVI
INTRODUCTION	may	7	0	0	7	0	2	17	0	2
	must	0	4	1	0	0	0	0	4	1
		law			med			com		
		EPI	DEO	EVI	EPI	DEO	EVI	EPI	DEO	EVI
CONCLUSION	may	6	0	0	1	0	0	0	0	1
	must	0	3	0	0	0	1	0	1	0

Table 4. Percentages of *may* and *must* (normalised data)

The results shown in Figure 1 reveal that there are no significant differences in epistemic uses of *may* in the introduction sections depending on the register domain accounting for 18.9% in computing texts, followed by 16.9% in legal texts and 15.9% in medical texts. In most of the instances, epistemic *may* indicates possibility and so they can be paraphrased by “it is possible that” (Palmer 1990: 41). Generic considerations may motivate this sort of interpretation: introductions are generally aimed at stating the authors’ attempt to solve a problem which, at such an early stage, tends to be done in terms of tentative possibility. The use of epistemic *may* in this section can be also taken as a mitigating device by means of which writers soften their positions so as to avoid imposing their views on their audience.

As for the expression of epistemicity in conclusions, the epistemic uses of *may* are predominant in the conclusion sections of legal texts. Its occurrence in the legal register accounts for 21% of the cases, followed by 3.8% in medical texts. The conclusions of computing texts do not contain epistemic uses of *may*. When compared to introductory sections, conclusions present lower frequencies of epistemic *may*. This is possibly due to the fact that conclusions, being the closing section of the research article, contain the solutions for the problems initially stated by the writers. In this final stage, solutions are likely to be presented either as feasible or infeasible on the basis of the empirical procedures which have been accounted for in the body of the article, allowing for lower levels of mitigation.

The expression of deonticity as shown by the use of *must* differs across registers: legal texts contain a noticeable number of instances of deontic *must* in introductions with a percentage of 9.7, followed by 4.4% in computing texts. The introductions of medical texts, on their part, are entirely lacking in deontic *must*. As for the conclusions, legal and computing texts present different levels of deonticity with percentages of 10.5 and 3.8, respectively. Higher levels of deonticity as found in legal texts are associated with the restrictions imposed by laws so that these uses of *must* tend to reveal legal obligations. Similar to what happens with their introductions, the conclusions of medical texts do not contain deontic uses of *must*. We contend that, in principle, writers keep to a minimum the expression of deonticity in the introductory sections of medical research articles as a way of avoiding to impose their ideas on the members of their discourse community. Although we have not registered the presence of deontic *must* in the conclusions of the medical research papers, it has been noted that the expression of deonticity in the conclusions of this type of articles does not appear to be associated with prescriptive attitudes since this genre is aimed at “present[ing] new knowledge rather than direct rules of action” (Vihla 1999: 22).

Finally, some evidential uses of modals have been attested in our corpus. Evidential *must* occurs in the introductions of legal texts in 2.4% of the cases and in the same section of computing texts in 1.1% of the cases. In conclusions, it has been only registered in medical texts in 3.8% of the cases. As for evidential *may*, the highest occurrence has been found in the introductions of medical texts in 4.5% of the

cases, followed by 3.8% in the conclusions of computing texts and 2.2% in the introductions of the articles of the same register.

4.1. INFERENTIAL EVIDENTIAL *MUST*

Must can convey meanings related to the performance of inferential processes by means of which individuals are led to some knowledge (Plungian 2001; Squartini 2008; Cornillie 2009). Even though scarce, evidential interpretations of this modal have been attested in our corpus. Let us consider the following example:

(4) Conversion to fixed-point representation and the resulting quantization effects has to be addressed. A forward pass through the sonar emulator neural network requires approximately 92000 multiply–accumulate operations. The FPGAs used in the SRC-6e can each perform 144 18-b multiplications in parallel [25]. The neural network *must* be mapped to this architecture. (Duren2007real-time-COM-INTRO)

Given the immediately preceding context where *must* appears in (4), the value attached to this modal is that of an inferential evidential marker since it indicates that the writer's knowledge as regards the neural network has been derived after an inferential process by which the expert reader would recover the intended meaning, i.e. the inferential mode of knowing. This is indicated by the employment of *the resulting quantization effects [...]*, *A forward pass through the sonar emulator neural network requires approximately [...]* and *The FPGAs used in the SRC-6e can each perform [...]*. The inferential process that takes place here is an example of reflected evidence since the writer has had a direct access to situations, mainly by empirical research, that allow her/him to arrive at a certain conclusion (*the neural network must be mapped to this architecture*) by means of the contextual evidences (previous sentences above mentioned).

4.2. INFERENTIAL EVIDENTIAL *MAY*

To our knowledge, inferential evidential uses of *may* have been previously identified in Alonso-Almeida (2010) and Alonso-Almeida and Cruz-García (2011) where its occurrence has been attested in the abstracts of scientific papers. In most cases, evidential uses of this modal indicate how the writer has been led to some knowledge. Examples (5) and (6) illustrate the inferential evidential value that can be attached to *may*:

(5) The learning algorithm developed in this paper is quite different from the traditional family of derivative-based descent methods. First, a constructive approach is used, which builds the network one node at a time. The advantages of a constructive approach include computational efficiency and the ease of determining a suitable network size. In fact, there is theoretical evidence to suggest that the learning problem *may* be intrinsically easier if we are allowed to add nodes and weights during the learning process [14]. Although constructive approaches are not guaranteed to produce networks of absolute minimal size, there is good reason to believe that they can produce representations which are efficient [15]. (Hush1998efficient-COM-INTRO)

(6) It can be concluded from the ILC's draft articles that State responsibility in international law arises when a breach of an international obligation occurs. It *may* therefore seem curious that the obligations contained with the principal treaties on unlawful interference in international civil aviation do not address the problem of terrorism directly. Rather the obligations contained therein are directed primarily at the response of other States to such acts of unlawful interference. (Young2003responsibility-LAW-CONC)

In (5) *may* can be taken as a manifestation of the writers' inferential reasoning, mainly deductive. This interpretation of the modal is further supported by its occurrence in the consequent of a conditional construction which explicitly indicates the inferential path to be followed. In both cases, the truth of the antecedent is a necessary but not a sufficient condition for the truth of the consequent and so there is a secondary intended meaning which is one of probability. As Cornillie (2009: 59) points out, the existence of primary evidential readings does not block the expression of epistemic commitment in different degrees. In terms of strength of assumptions, the degree of confidence with which the assumption in which *may* is embedded is by no means the highest and, consequently, cannot be considered to be certain. Ideally, after optimal processing of these fragments, the first interpretation to occur to the addressee is that the existing relation between the conjuncts of the conditional is that of compatibility rather than entailment.

As in the previous examples, context needs to be enriched in various ways in sample (6) so as to arrive at the inferential evidential interpretation of *may*, which in this case appears embedded in a construction of the form *it seems + adjective + that + clause* which has been labelled as multiple qualification (Johansson 2001: 240). In this example, *may* marks the possibility that, for someone knowing the relationship between State responsibility in international law and acts of unlawful interference in international civil aviation, it is possible that s/he finds it unusual that the problem of terrorism is not addressed in the ILC's draft articles in a straightforward way. The inferential interpretation of *may* in this example is strengthened by *It can be*

concluded [...], and the adverbial particle *therefore* indicating result of a logical process. Moreover, the use of the verb *seem* is worth mentioning here.

Seem has been taken to possess a variety of meanings depending on the syntactic construction in which it occurs. Aijmer points out that “*seem* straddles the boundary between perception (an evidential category) and epistemic modality” (Aijmer 2009: 64). She defines *seem* as a phenomenon-based perception verb the meaning of which is related to the existence of an external stimulus which originates a reaction in the subject such as a belief that something may be the case.

In our example, *seem* also indicates inferential reasoning on the basis of the information made manifest earlier in the research paper about the relationship that holds between unlawful interference and terrorism. Although initially *seem* can be taken as a visual evidential which indicates that the information has been acquired through the senses, mainly seeing, i.e. the communicator has accessed the information in the ILC’s draft articles by reading them. *Seem* can ultimately be taken as a cognitive evidential which points at the information’s being acquired inferentially, i.e. inferential processes operate on the information firstly acquired through the senses. Aijmer (2009: 72) considers it as a perception verb and an evidential verb of cognition and Alonso-Almeida and Lareo (forthcoming) add that by using this kind of constructions (*Seem* + *that* clause) “the speaker reaches this information by cognitively processing perception inputs and his (shared) knowledge of the world”. A pragmatic effect derived from the use of this evidential verb in a construction where the experiencer has been elided in favour of the impersonal pronoun *it* is that of lowering the degree of speaker commitment and/or certainty. In this sense, *may* can also signal epistemicity.

5. CONCLUSION

In this paper we have focused on the use of the modals *may* and *must* as indicators of inferentially acquired knowledge in a corpus of scientific computing, legal and medical research papers. Even though evidential interpretations of these modals have proved to be less frequent than epistemic or deontic ones, they have been attested in our corpus. Differences in the expression of deonticity, epistemicity and evidentiality seem to be subject to register and genre variation: epistemic interpretations of *may* abound in the introduction sections of all the registers analysed. This is so because in the introductory sections to research papers the authors’ intention to solve a certain problem is given in terms of tentative possibility.

The occurrence of *must* with deontic value is register and genre dependent: both legal and computing texts tend to favour the expression of deonticity, mainly obligation, in their introductions and conclusions, while medical texts do not so in neither of these two sections. The prevalence of prescriptive attitudes associated with deontic *must* in legal texts, for instance, may be derived from the restrictions imposed

by laws, while its absence in the introductions of medical texts may be associated with the writers' need to avoid sounding authoritative to their peers from the outset.

As for the expression of evidentiality, we have shown that the English modals *may* and *must* are indeed used to indicate mode of knowing, mainly inferential. *May* seems to be associated to the expression of inferential mode of knowing in a more explicit way since it appears in the conclusions of logical processes constraining the assumptions to be recovered during the interpretation process in which case *may* presents both evidential and epistemic readings.

Must, on its part, occurs in propositions where the conclusions of logical processes are expressed. In both cases, we have shown that appropriate contextual enrichment allows the recovery of the intended meaning, i.e. mode of knowing. However, quantitative results point at the need of widening the corpus of study. Future research may focus on the study of evidentiality markers, not only as indicators of source of knowledge and/or mode of knowing, but also as indexical of the author's stance, in particular, in relation to his/her subjective/intersubjective positioning.

REFERENCES

- Aikhenvald, Alexandra (2004). *Evidentiality*. Oxford: Oxford University Press.
- Aijmer, Karin (1980). *Evidence and the declarative sentence*. Stockholm: Almqvist-Wiksell.
- Aijmer, Karin, ed. (2001). *A wealth of English. Studies in honour of Göran Kjellmer* (Gothenburg Studies in English 81). Göteborg: Acta Universitatis Gothoburgensis.
- Aijmer, Karin (2009). Seem and evidentiality. *Functions of Language* 16.1: 63-88.
- Alonso-Almeida, Francisco (2010). Evidential *may* in medical scientific abstracts. In Moskowich-Spiegel-Fandiño, Isabel, Crespo-García, Begoña and Paula Lojo-Sandino, eds. 29-37.
- Alonso-Almeida, Francisco and Laura Cruz-García (2011). The value of *may* as an evidential and epistemic marker in English medical abstracts. *Studia Anglica Posnaniensia* 46.3: 59-73.
- Alonso-Almeida, Francisco and Inés Lareo (forthcoming). The status of *seem* in nineteenth-century CePhit.
- Anthony, Laurence (1999). Writing research article introductions in software engineering: How accurate is a standard model? *IEEE Transactions on Professional Communication* 42.1: 38-46.
- Boye, Kasper (2010). Semantic maps and the identification of cross-linguistic generic categories: Evidentiality and its relation to epistemic modality. *Linguistic Discovery* 8.1: 4-22.

- Bybee, Joan (1985). *Morphology: A study of the relation between meaning and form* (Typological Studies in Language 9). Amsterdam: John Benjamins.
- Carretero-Lapeyre, Marta (2004). The role of evidentiality and epistemic modality in three English spoken texts from legal proceedings. In Marín-Arrese, Juana, ed., 25-62. Chafe, Wallace (1986). Evidentiality in English conversation and academic writing. In Chafe, Wallace and Johanna Nichols, eds., 261-272.
- Chafe, Wallace and Johanna Nichols, eds. (1986). *Evidentiality: The linguistic coding of epistemology*. New Jersey: Ablex.
- Cooper, Charles (1985). *Aspects of article introductions in IEEE publications*. Aston: Aston University MSc Dissertation.
- Cornillie, Bert (2009). Evidentiality and epistemic modality. On the close relationship between two different categories. *Functions of Language* 16.1: 44-62.
- Coulthard, Malcolm, eds. (1986). *Talking about text*. Birmingham: English Language Research.
- Coulthard, Malcolm, eds. (1994). *Advances in written text analysis*. London: Routledge.
- Crookes, Graham (1986). Towards a validated analysis of scientific text structure. *Applied Linguistics* 7: 57-70.
- De Haan, Ferdinand (1999). Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics* 18: 83-101.
- De Haan, Ferdinand (2006). Typological approaches to modality. In Frawley, William, Eschenroeder, Erin, Mills, Sarah and Thao Nguyen, eds., 27-69.
- Dendale, Patrick and Liliane Tasmowski (2001). Introduction: Evidentiality and related notions. *Journal of Pragmatics* 33: 339-348.
- Downing, Angela and Phillip Locke (1992). *A university course in English grammar*. Hemel Hempstead: Phonix ELT.
- Dudley-Evans, Tony (1986). Genre analysis: An investigation of the introduction and discussion sections of MSc dissertations. In Coulthard, Malcolm, eds., 128-145.
- Dudley-Evans, Tony (1994). Genre analysis: An approach to text analysis for ESP. In Coulthard, Malcolm, ed. 219-228.
- Dudley-Evans, Tony (2000). Genre analysis: A key to a theory of ESP? *IBÉRICA* 2: 3-11.
- Frawley, William, Eschenroeder, Erin, Mills, Sarah and Thao Nguyen, eds. (2006). *Typological approaches to modality*. Berlin: Mouton de Gruyter.
- Givón, Talmy (1982). Evidentiality and epistemic space. *Studies in Language* 6: 23-49.
- Ifantidou, Elly (2001). *Evidentials and relevance*. Amsterdam: Benjamins.
- Johansson, Stig (2001). The English verb seem and its correspondences in Norwegian: What seems to be the problem? In Aijmer, Karin, ed., 221-245.
- Kastovsky, Dieter and Aleksander Szwedek, eds. (1986). *Linguistics across historical and geographical boundaries. In honour of Jacek Fisiak on the occasion of his*

- 50 birthday (Vol. 2: *Descriptive, Contrastive and Applied Linguistics*). Berlin: Mouton de Gruyter.
- Klinge, Alex (1993). The English modal auxiliaries: From lexical semantics to utterance interpretation. *Journal of Linguistics* 29: 315-357.
- Lyons, John (1977). *Semantics*. Cambridge: Cambridge University Press.
- Marin-Arrese, Juana, ed. (2004). *Perspectives on evidentiality and epistemic modality*. Madrid: Editorial Complutense.
- Mayer, Rolf (1990). Abstraction, context and perspectivization – Evidentials in discourse semantics. *Theoretical Linguistics* 16: 101-163.
- McCready, Eric and Norry Ogata (2007). Evidentiality, modality and probability. *Linguistics and Philosophy* 30: 147-206.
- Mithun, Marianne (1986). Evidential diachrony in Northern Iroquian. In Chafe, Wallace and Johanna Nichols, eds., 89-112.
- Moskowich-Spiegel-Fandiño, Isabel, Crespo-García, Begoña and Paula Lojo-Sandino, eds. (2010). *Language windowing through corpora/Visualización del lenguaje a través de corpus*. A Coruña: Universidade da Coruña.
- Ngozi-Nwogu, Kevin (1997). The medical research paper: Structure and functions. *English for Specific Purposes* 16.2: 119-138.
- Nuyts, Jan (2001). *Epistemic modality, language and conceptualization*. Amsterdam: Benjamins.
- Palmer, Frank (1986). *Mood and modality*. Cambridge: Cambridge University Press.
- Palmer, Frank (1990). *Modality and the English modals* (2nd ed.). London: Longman.
- Palmer, Frank (2001). *Mood and modality* (2nd ed.). Cambridge: Cambridge University Press.
- Papafragou, Anna (2000). *Modality. Issues in the semantics-pragmatics interface*. Oxford: Elsevier.
- Plungian, Vladimir (2001). The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33: 349-357.
- Posteguillo, Santiago (1999). The schematic structure of computer science research articles. *English for Specific Purposes* 18.2: 139-160.
- Radden, Günter and René Dirven (2007). *Cognitive English grammar* (Cognitive Linguistics in Practice (CLiP)). Amsterdam and Philadelphia: John Benjamins.
- Ruiying, Yang and Desmond Allison (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes* 22: 365-385.
- Samraj, Betty (2002). Introductions in research articles: Variation across disciplines. *English for Specific Purposes* 21: 1-17.
- Squartini, Mario (2008). Lexical vs. grammatical evidentiality in French and Italian. *Linguistics* 46.5: 917-947.
- Swales, John (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press.

- Van der Auwera, Johan (1986). The possibilities of *may* and *can*. In Kastovsky, Dieter and Aleksander Szwedek, eds., 1067-1076.
- Van der Auwera, Johan and Vladimir Plungian (1998). On modality's semantic map. *Linguistic Typology* 2: 79-124.
- Vihla, Minna (1999). *Medical writing. Modality in focus*. Amsterdam: Rodopi.
- Willet, Thomas (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language* 12: 51-97.

APPENDIX

In this appendix, we include the journals used to carry out this study per register domain (computing, law, and medicine):

COMPUTING

ACM Computing Surveys
ACM Transactions on Graphics
IEEE Transactions on Evolutionary Computation
IEEE Transactions on Image Processing
IEEE Transactions on Neural Networks
IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on Software Engineering
IEEE Transactions on Wireless Communications
IEEE/ACM Transactions on Networking
International Journal of Computer Vision

LAW

Air & Space Law/Supplement Air & Space Law
American Journal of Criminal Law
Business Law Review
Family Court Review
International Journal of Law and Psychiatry
International Journal of Law in Context
International Journal of the Sociology of Law
Legal Studies
Legal Theory
The International Journal of Comparative Labour Law and Industrial Relations

MEDICINE

Academic Emergency Medicine
Aesthetic Plastic Surgery
Basic Applied Myology
BMC Pulmonary Medicine
Breast Cancer Research
Harm Reduction Journal
Journal of Laboratory and Clinical Medicine
Journal of Urban Health: Bulletin of the New York Academy of Medicine
The AAPS Journal
The New England Journal of Medicine