EPISTEMIC MODALS IN EARLY MODERN ENGLISH HISTORY TEXTS. ANALYSIS OF GENDER VARIATION

Francisco J. Álvarez Gil

Universidad de Las Palmas de Gran Canaria

Abstract: In the present study, samples of late Modern English scientific texts have been analysed to evaluate cases of epistemic modality as realised by modal verbs. The aim of this research was to detect if there exist variances in the way modals are used in historical texts from a gender perspective. For this, I have interrogated the Corpus History English Texts (1700-1900) which is part of The Coruña Corpus of English Scientific Texts, which contains history texts written by male and female authors. I have used the Coruña Corpus Tool for retrieval, although manual analyses have been carried out as well. Each of the occurrences found have been categorised according to their contextual meanings. The results obtained account for a high frequency on the usage of these modal verbs according to gender and the diverse pragmatic functions these modal verbs accomplish in the communicative process, such as mitigation and negative politeness. From a pragmatic perspective, epistemic modals have the potential to allow negotiation of meaning between writers and their audience among other functions.

Keywords: epistemic meaning; modality; modal verbs; early Modern English

1. INTRODUCTION

The research conducted has focused on samples from English scientific texts from 1700 to 1900 in order to evaluate epistemic modality as realised by modal verbs. Epistemic modality seems to be strongly connected to the idea of *truth* and the authors' responsibility and commitment regarding their statements (Traugott, 1989; Sweetser, 1990; Stukker Sanders and Verhagen, 2009). A related notion to epistemic modality is evidentiality. For some scholars evidentiality represents a subdomain of epistemic modality. Other scholars consider evidentiality as an independent category. Depending on how these concepts relate, these two concepts are divided into disjunction, inclusion, and intersection (Dendale and Tasmowski, 2001). I shall follow the disjunctive approach in this paper in line with Cornillie (2009) who argues that the mode of knowing should not be associated with the degree of the authors' commitment towards their texts. I will come to this again in section two.

My interest was to see whether differences in the use of these modals could be detected from a gender perspective. For this, I have interrogated the subcorpus of History of *The Coruña Corpus of English Scientific Writing* (CC), which contains representative extracts of English historical texts written between 1700 and 1900, using its own retrieval tool, i.e. the Coruña Corpus Tool. Each occurrence has been categorised according to its contextual meaning following Dixon's description of modal verbs that claims there are modals and what I can call semi-modals, which express the modalities (2009: 172). However, there are also other valuable insightful studies on modals as Coates (1983), Leech (1971) and Palmer (1979), among others, which have served as references for the present study.

The process followed in my analysis consists basically in the following: firstly, I have produced a list of occurrences in the corpus to check the presence of modal verbs in the history texts available. Secondly, I have interrogated and analysed the corpus to find the pragmatic functions those modals play in the different texts. Finally, I have checked the results to find out if there exists any difference in the use of epistemic modals in late

Modern English scientific register regarding the gender of the writers. For this paper, I have chosen modal verbs *can, may, might,* and *must*, as they are frequently associated with this modal meaning.

Results report on frequency of usage of these modal verbs according to gender, but, most importantly, the different pragmatic functions these modal verbs fulfil in the communicative process. One such pragmatic function is mitigation of claims (Alonso Almeida, 2015), and so modals are used as a negative politeness strategy (Brown and Levinson, 1987) in order to avoid or minimize imposition, to hedge the illocutionary force of a specific statement, or to put social distance so as to save the author's face. In this sense, modals are quite useful as they enable an interactive construction of scientific knowledge giving the chance to the writer and the readers to negotiate meaning.

2. MODALS, EPISTEMIC MODALITY AND RELATED NOTIONS

The study of modal verbs in the history of English is quite controversial. Scholars do not seem to agree as for the exact dating of this innovation in the language, and whether it is an innovation, or it existed already in OE. The way in which the issue is methodologically approached may change our view of the facts. Whereas Lightfoot (1979) and Roberts (1985) offer a syntactic explanation as to the development of this closed class of verbs, Plank (1983) considers grammaticalization as the correct path to describe the emergence of modals. Plank (1983) is less in the line of the generative Principle of Transparency posed by Lightfoot (1979). Other scholars, however, seek to reconcile the methods combining aspects of formal syntactic theory to explain syntactic changes within a more functional and inclusive methodology.

In PDE, there are nine central modal verbs, according to Biber *et al.* (1999:483): *can could, may, might, shall, should, will, would, must.* A further group of modals is the peripheral modals or marginal modals, also called semi-modals: *need (to), ought to, dare (to), used to*. This group is characterised because they tend to take *to*-infinitive rather than bare infinitive, but they share other features with modals such as direct negation with *not* (also in contracted form), and inversion in questions (Biber *et al.,* 1999). However, verbs like *dare* and *need* also take on occasions the periphrastic *do*.

These same authors make a correlation of pair modals according to tense: *can-could, may-might, shall-should, will-would* (Biber *et al.,* 1999:485). However, they admit some of the forms used to express a past event may also appear in contexts which show the speaker's stance towards the proposition manifested and the expression of hypothetical situations, conveying tentativeness and politeness, rather than past tense, as in *Could I sit here a minute, Joyce?* For this, they 'regard modal verbs as unmarked for tense'.

Modal verbs are but one way in which modality can be linguistically accomplished. There are varied approaches to the notion of modality, but they all share the fact that this notion owns a cognitive dimension, which includes the speaker's viewpoints and attitudes towards a given propositional content (Lyon, 1977:452; Bybee *et al.,* 1994; Palmer, 1986:2). What stands water clear is that modality is not a uniform monolithic concept. The representation of matters of opinions, beliefs and attitudes suggests an array of alternative modality taxonomies depending on the criteria for these classifications.

Epistemic modality has been defined as the evaluation of chances concerning the actualization of an event (Nuyts, 2001:21), and so it seems to be strongly connected to the idea of *truth* and the authors' responsibility concerning their statements (Stukker *et al.,* 2009; Traugott, 1989). Epistemic modality markers and devices are defined as "linguistic expressions that qualify the truth value of a propositional content" (Vold, 2006:65), and these can be modals (*may, might*), adverbials (*possibly, probably*), and matrices (*It is possible that…*).

In the case of the relationship between epistemic modality and evidentiality, this could be of three main types: inclusive, intersective and disjunctive. The inclusive type considers epistemic modality to be a subcategory of epistemic modality, and so evidentials evaluate the chances one proposition (*P*) has to be true. The intersective approach refers to a sort of continuum between degrees of certainty concerning the actualization of *P* based on the evaluation of the evidences a speaker has to claim *P*. In this line, Van der Auwera and Plungian (1998) highlight the notion of evidential inferentiality, which advocates the intersective relationship between epistemic necessity and inference. One clear example is *That must be the postman*. The epistemic modal verb *must* indicates the speaker's inference in the light of the contextual premises selected. This intersective view is supported in Diewald *et al.* (2009:190): "Evidentiality is concerned with indicating the information source the speaker is relying on to make a claim. This places this category next to epistemic modality without, however, merging them into one".

There is yet another type, i.e. the disjunctive, which considers evidentiality a category on its own. This notion of evidentiality is followed in Cornillie (2009) and Alonso-Almeida (2015). The idea behind is that evidentials inform on the speaker's involvement to claim *P* rather than on the evaluation of chances of *P* to be true. Normally, the

assessment of *P* in terms of truth is performed on the evaluation of the evidential. In this sense, following Willet (1988), first hand (visual, auditory) evidence is said to be more reliable than third hand (hearsay, reported) evidence and inferential meaning. This view is challenged in Alonso-Almeida (2015), and this author claims that evaluation of source of knowledge does not necessarily entail different degrees of propositional truth.

3. CORPUS AND METHODOLOGY

3.1. The Corpus of English History Texts (CHET)

The corpus used for the present research has been The *Corpus of English History Texts* (henceforth CHET), which is one of the sub-corpuses within the *Coruña Corpus of English Scientific Writing*, and it aims at compiling Modern English, history texts. Crespo and Moskowich (2010) have offered a substantial description of this corpus in their paper *A Corpus of History English Texts (CHET) as Part of the Coruña Corpus Project,* description that is followed in the present work.

In this sense, several scientific landmarks have been considered in order to limit the time-span represented in the sub-corpus of our choice. The first text in CHET dates back to 1704. The end of the sixteenth century and the beginning of the seventeenth century have been identified by Taavitsainen and Pahta (1997) as the moment at which the medieval scholastic thought-style started to be gradually superseded by new patterns of thought and new methodological procedures based on observation started to be used. The foundation of the Royal Society in 1660 and the publication of the guidelines for presenting scientific works in a clear and simple way had a greater impact on accentuating the importance of style in scientific communication.

The last text in CHET dates back to 1895. Again, the end of the nineteenth century and the beginning of the twentieth century roughly coincide with some important events in the history of science such as the discovery of the electron (1896), the formulation of Planck's *Quantum Theory* (1900) and the publication of Einstein's *Special Theory of Relativity* (1905). Obviously, all of these events brought about the need to change scientific discourse patterns, as put forward by Huxley in the 1897 *International Congress of Mathematics*. As regards the genres represented in CHET, there are articles, essays, lectures, textbooks and treatises written by both male and female authors. The distribution of texts in CHET regarding genres is shown in figure 1.
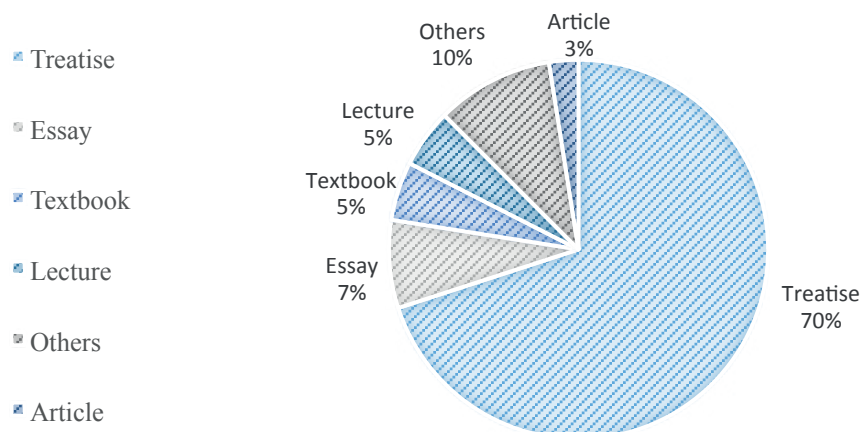

Figure 1. Distribution of words per genre in *CHET.*

As to the size of CHET, it covers about 400,000 words. The relevance of this data lies in the fact that previous studies have shown that 1,000-word samples are not really enough for the study of variation within the scientific register (Biber, 1993) mainly because the scientific register was not as standardized at that time as it is nowadays (Crespo and Moskowich, 2010). For the present study, it is relevant to specify the distribution of words according to the authors' sex. In CHET there are 81,775 words written by female authors and 322,842 by male authors.

Table 1. *Words in CHET* (after Crespo and Moskowich 2015).

Eighteenth century	201,794
Nineteenth century	202,823
Total	404,617

3.2. Methodology

In relation to the methodology used, I have interrogated the subcorpus of History of *The Coruña Corpus of English Scientific Writing*, using its own retrieval tool, i.e. the Coruña Corpus Tool. My interest was to see whether differences in the use of these modals could be detected from a gender perspective. The process followed consists basically in the following: firstly, I have produced a list of occurrences in the corpus to check the presence of modal verbs in the history texts available. Secondly, I have interrogated and analysed the corpus to find the pragmatic functions those modals play in the different texts. Finally, I have checked the results to find out if there exists any difference in the use of epistemic modals in late Modern English scientific register regarding the gender of the writers.

4. RESULTS

The results of my enquiries are given, below, in Table 2, and also shown in the accompanying Graph 1.

Table 2. Results and variation statistics.

	CAN				MAY			
	Male Author	Female Author	*p > 0.05*	*LL*	Male Author	Female Author	*p > 0.05*	*LL*
Possibility	0.00650	0.00489	0.06011	0.29	0.05080	0.04280	0.08000	0.88
Probability	-	-			0.00960	0.00489	0.00471	1.91
Epistemic necessity	-	-			-	-		

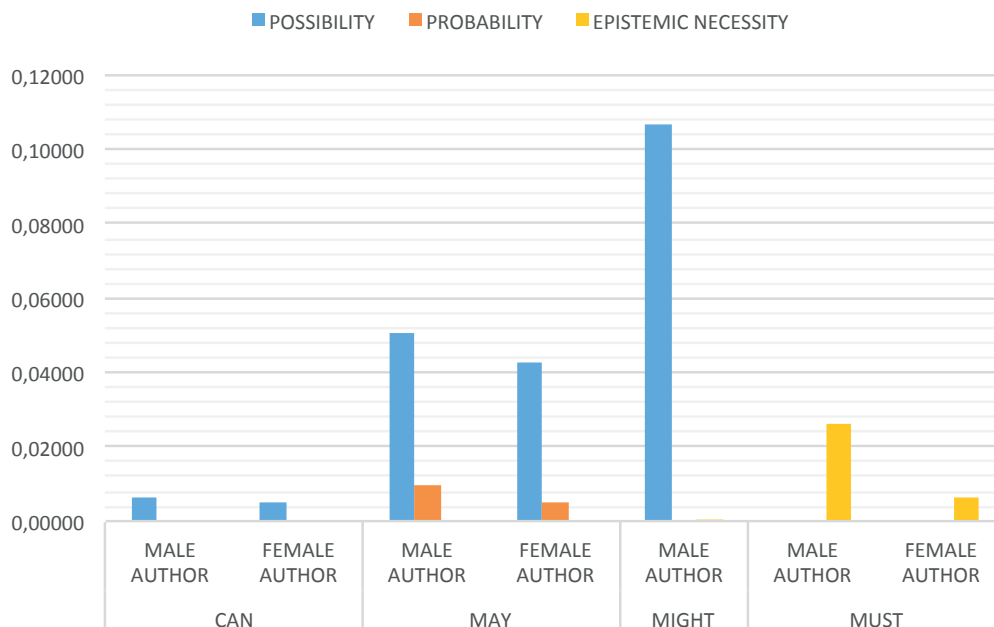|	MIGHT				MUST			
	Male Author	Female Author	*p > 0.05*	LL	Male Author	Female Author	*p > 0.05*	LL
Possibility	0.10686	0.10394	0.00292	F	-	-		
Probability	-	-			-	-		
Epistemic necessity	0.00031	0.01467	0.01436	F	0.02633	0.00611	0.02022	15.75

Figure 2. Distribution of epistemic modal verbs *can, may, might, must* in CHET, gendered.

Data have been normalized taking into account the number of words written by men and women in order to make them comparable. In addition, I have calculated the log likelihood ratio and the frequency variation to see whether this variation is significant. The results have been organized according to meanings. The possibility meaning refers to the epistemic domain. The dynamic sense appears extensively in the two subcorpora with no significant variation from a gender perspective. The following are examples of this type of modal. Note that, in the case of (4), *can* is used to indicate reliance in one's cognitive abilities to meet a particular conclusion. In other words, it indicates disposition. In all the examples, M stands for *male* and F stands for *female*.

(1) Water-men by Sea: They **can** make their Kings deliver Glorious Speeches, which were never ſpoken; and can dub Princes and give them Titles, that never had any Being, but under a ſhaven Pate (hist 1705 Anderson 48-103. txt) M

(2) The common *Men* **can** go through great Hardſhips; and will live in Places, and on ſuch mean Food, that would kill our Natives (hist 1739 Justice 1-46.txt) F

(3) As to the Ruſſia Climate, it is extraordinary cold, as you **may** imagine: For I ſaw Two and Thirty Thouſand Men exerciſed upon the Ice. (hist 1739 Justice 1-46.txt) F

Epistemic possibility is well represented in the corpus analysed, and this is realised by *can*, *may*, and *might,* which I exemplify below:

(4) William Biſhop of Durham, who **can** be no other, but he who is commonly called, Willelmus de Carilefo; and likewiſe he and Turgot Prior of Durham, are placed among the Witneſſes in the two ſuppoſed Charters of our Edgar, inſinuating Homage (hist 1705 Anderson 48-103.txt) M

(5) For her Entertainment, there are twice a Week Italian Opera's, which are at Her Majeſty's Expence; and none **can** be admitted but thoſe who have a Ticket (hist 1739 Justice 1-46.txt) F

(6) And however theſe things **may** to ſome ſeem dry and Barren, yet even theſe Seals protect us from Forgery (hist 1705 Anderson 48-103.txt) M

(7) These with Cîteaux herself **may** possibly make up the ten mentioned, says Manriquez, in the collection of "diffinitiones" of 1134 (hist 1893 Cooke 625-648.txt) F

(8) Some of the Indians fled, some seized their arms, but La Salle, alone and unattended, was in the midst, but he did not present the calumet, lest it **might** be regarded as an evidence of conscious weakness (hist 1884 Breese 67-118.txt) M

(9) that the war with Sparta **might** be carried on vigorously (hist 1857 Sewell 226-260.txt) F

Of the three modal verbs, *might* is more frequently used than the other two. There is however no significant variation according to gender. The modal verb *may* does not show significant variation, either. The modal verb *can* does indicate some variation, as this modal is more frequently used with an epistemic sense by men than by women with a log likelihood ratio of 0.29 with overuse in the case of the subcorpus of texts written by male scientists and a *p* value of 0.06011. The epistemic sense of these forms is sometimes strengthened by the use of particles, such as *possibly* in (7), or by embedding them in conditional structures, as in (8).

Epistemic probability is poorly represented in this corpus, and this meaning is given in some instances of *may*. The distribution in the two subcorpora is not significant. In raw numbers, I have detected 31 cases of this meaning in the case of texts written by men and 4 cases in those written by women. Examples are the following:

(10) The words noſtra ripa **may** probably fignify the Roman or ſouth ſide of the river or aeſtuary of Tay. who told the whole adventure (hist 1732 Horsley 38-55.txt) M

(11) It **may** probably be placed between the years 1151 and 1154, as the charter of Henry, which might, however, it is to be remembered, either precede or follow the foundation, is given as duke of Normandy and count of Anjou (hist 1893 Cooke 625-648.txt) F

In both cases, the modal verb is accompanied by the hedge *probably*, whose function is to contribute to mitigate the propositional content. The combination of these two epistemic devices seems to follow from the authors' intention to avoid imposition.

Epistemic necessity is registered in some instances of *might* and *must* in both subcorpora. While the form *might* is the preferred form by female writers (*p* 0.01467), *must* is extensively used to convey epistemic necessity by male writers (*p* 0.02022), and so it is statistically significant in both cases. Examples are the following:

(12) Stephen, aware of this circumftance, and apprehending that, while he was engaged in the fiege, his enemies **might** be making a dangerous progrefs under the conduct of the Earl of Gloucefter, whofe vigour and activity experience had taught him to dread, might juftly deem it an act of policy to permit the emprefs to join her brother, as his enemies would then be concentrated in one point, againft which he would be enabled to direct his whole force (hist 1790 Gifford 179-189.txt) M

(13) The Athenians had no money, no ships, no soldiers, —one **might** have supposed that their spirits must have sunk completely. But they did not. (hist 1857 Sewell 226-260.txt) F

(14) I fhall here add a Remarkable Tranfaction, which **muft** have been done about this time (if it were ever done at all) which I have fome reafon to doubt it was not, becaufe our own Hiftorians are wholy filent in it (hist 1704 Tyrrell 952-968.txt) M

(15) At this crifis, had general Gage ventured without his entrenchments, both the American army and the people, **muft** have been involved in extreme diftrefs (hist 1805 Warren 229-277.txt) F

All the examples quoted above show cases of complex modal structures with the modals followed by the progressive in (12) and the perfective in the other three cases. This is what Boye and Harder (2009) call evidential substance. This structure seems to fit very well in history texts, as it reflects disciplinary tradition. Historians, lacking some first-hand knowledge, are able to make inferences in the light of the evidences they have. This cognitive material is still matter of future criticism in the event new evidences appear. The difference in the use of *must* and *might* from a gender perspective can be accounted for in terms of politeness. The use of *might* as the *irrealis* of *may* (Palmer, 2001) in the case of female authors may suggest a marked indication of avoidance of imposition.

5. CONCLUSION

This paper has shown that there are differences between men and women in the use of modal verbs with an epistemic meaning, but most importantly evidentiary inferentiality. In the want of more research, my findings seem to be statistically significant. I have detected a massive use of *may* and *might* to designate epistemic possibility in the scientific texts analysed. One small proportion of *might* is devoted to entail probability, which refers also to an epistemic sense. Finally, I have also registered the use of *must* and *might* to communicate inferentiality. In the sense described here in section 2, these two forms are evidentials, as they suggest authorial involvement in the elaboration of meaning. In the examples under scrutiny, these verbs show the cognitive source of the knowledge presented, as conclusions develop from the author's own critical thinking in the light of the evidences at hand.

More research should be done to cover more texts from other subcorpuses of the Coruña Corpus to see whether variation identified in CHET concerning inferential *must* and *might* is idiosyncratic, or it represents a contemporary tendency in the scientific register.

ACKNOWLEDGMENTS

REFERENCES

Alonso-Almeida, F. and L. Cruz-García. (2011). "The value of *may* as an evidential and epistemic marker in English medical abstracts". *Studia Anglica Posnaniensia, 46*/3: 59-73. https://doi.org/10.2478/v10121-010-0004-7

Alonso-Almeida, F. (2015). "On the mitigating function of modality and evidentiality from English and Spanish medical research papers". *Intercultural Pragmatics, 12*/1: 33-57. https://doi.org/10.1515/ip-2015-0002

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing, 19*: 219-241. https://doi.org/10.1093/llc/8.4.243

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education Ltd.

Boye, K. and Harder, P. (2009). "Evidentiality: linguistic categories and grammaticalization". *Functions of Language, 16*: 9-43. https://doi.org/10.1075/fol.16.1.03boy

Brown, P. and Levinson, S.C. (1987). *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press.

Bybee, J., Revere P. and William P. (1994). *The evolution of grammar. Tense, aspect, and modality in the languages of the world*. Chicago/London: The University of Chicago Press.

Carretero, M. (2004). "The role of evidentiality and epistemic modality in three English spoken texts from legal proceedings". *Perspectives on Evidentiality and Epistemic Modality*. Ed. J. I. Marín-Arrese. Madrid: Editorial Complutense. 25-62.

Chafe, W. (1986). "Evidentiality in English conversation and academic writing". *Evidentiality: The Linguistic Coding of Epistemology*. Eds. W. Chafe and J. Nichols Norwood. NJ: Ablex. 261-272.

Coates, J. (1983). *The semantics of modal auxiliaries*. London/Canberra: Croom Helm.

Cornillie, B. (2009). "Evidentiality and epistemic modality. On the close relationship between two different categories". *Functions of Language, 16*/1: 44-62. https://doi.org/10.1075/fol.16.1.04cor

Crespo, B. and Moskowich, I. (2015). "A Corpus of History Texts (CHET) as part of the Coruña Corpus Project" in *Proceedings of the international scientific conference*. St. Petersburgh State University, 14-23.

Dendale, P. and Tasmowski, L. (2001). "Introduction: Evidentiality and related notions". *Journal of Pragmatics, 33*: 339-348. https://doi.org/10.1016/S0378-2166(00)00005-9

Diewald, G., Kresic, M. and Smirnova, E. (2009). "The grammaticalization channels of evidentials and modal particles in German: Integration in textual structures as a common feature" in Hansen, M.M. and Visconti J. (eds.) *Current Trends in Diachronic Semantics and Pragmatics*. UK: Emerald, 189-209. https://doi.org/10.1163/9789004253216_011

Dixon, R. W. (2005). *A Semantic Approach to English Grammar.* Oxford: Oxford University Press.

Hunston, S. (2007). "Using a corpus to investigate stance quantitatively and qualitatively". *Stancetaking in Discourse*. Ed. Robert Englebretson. Amsterdam/Philadelphia: John Benjamins. 27-48. https://doi.org/10.1075/pbns.164.03hun

Hunston, S. and Thompson, G. (2000). *Evaluation in Text.* Oxford: Oxford University Press.

Hyland, K. (1998). *Hedging in Scientific Research Articles*. Amsterdam/ Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.54

Hyland, K. (2001). "Bringing in the reader: Addressee features in academic article". *Written Communication, 18*/4: 549-574. https://doi.org/10.1177/0741088301018004005

Leech, G. (1974). *Semantics*. Harmondsworth: Penguin Books.

Lightfoot, D. 1979. "Rule Classes and Syntactic Change". *Linguistic Inquiry, 10*: 83-108. London: Cambridge University Press.

Lyons, J. (1977). *Semantics*. Volume 2. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165693

Nuyts, J. (2001). *Epistemic modality, language and conceptualization. A cognitive-pragmatic perspective*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/hcp.5

Palmer, F. R. (2001). *Modality and the English Modals*. London: Longman.

Palmer, F. R. (1986). *Mood and modality* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Palmer, F. R. (1979). *Modality and the English modals*. London: Longman.

Perkins, M. (1983). *Modal expressions in English*. London: Frances Pinter.

Plank, F. (1983). Coming into Being among the Anglo-Saxons. In Davenport, Michael; Hansen, Erik and Nielsen, Hans Frede (eds). *Current Topics in English Historical Linguistics*. Odense: Odense University Press, 239-278.

Roberts, I. (1985). "Agreement parameters and the development of English modal auxiliaries". *Natural Language and Linguistic Theory, 3*: 21-58. https://doi.org/10.1007/BF00205413

Salager-Meyer, F. 1992. "A text–type and move analysis study of verb tense and modality distribution in medical English abstracts", *English for Specific Purposes, 11*/2: 93-113. https://doi.org/10.1016/S0889-4906(05)80002-X

Salkie, R. (1996). "Modality in English and French: A corpus–based approach", *Language Sciences, 18*/1-2: 381–392. https://doi.org/10.1016/0388-0001(96)00026-5

Stukker, N., Sanders,T. and Verhagen, A. (2009). "Categories of Subjectivity in Dutch Causal Connectives: A Usage-Based Analysis". *Causal Categories in Discourse and Cognition*. Eds. T. Sanders and E. Sweetser. Mouton de Gruyter: Berlin. 119-172. https://doi.org/10.1515/9783110224429.119

Sweetser, E. (1990). *From etymology to pragmatics. Metaphoric and cultural aspects of semantic structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620904

Taavitsainen, I. and Päivi P. (1997). Corpus of early English medical writing 1375-1550. *ICAME Journal, 21*: 71-78.

Traugott, E. C. (1989). "On the rise of epistemic meanings in English: An example of subjectification in semantic change". *Language, 57*: 33-65. https://doi.org/10.2307/414841

Van der Auwera, J. and Plungian, V.A. (1998). "Modality's semantic map". *Linguistic Typology, 2*: 79-124. https://doi.org/10.1515/lity.1998.2.1.79

Vold, E. (2006). "Epistemic modality markers in research articles: A cross-linguistic and cross-disciplinary". *International Journal of Applied Linguistics, 16*: 61-87. https://doi.org/10.1111/j.1473-4192.2006.00106.x

Willett, T. (1988). "A Cross-linguistic Survey of Grammaticalization of Evidentiality", *Studies in Language, 12*: 51-97. https://doi.org/10.1075/sl.12.1.04wil